

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2004, 07:40:11 ; Search time 2319.58 Seconds
(without alignments)
9398.725 Million cell updates/sec

Title: US-09-864-675-3
Perfect score: 897
Sequence: 1 atgaggcgcgacccggcccc.....caatgggtcaacttctcctaa 897

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*

```
29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c	1	674	75.1	805	12	BI918620
	2	565.6	63.1	1047	12	BM914622
	3	467.2	52.1	1041	12	BI412864
	4	467	52.1	524	13	BX281777
	5	408.2	45.5	549	9	AA706226
	6	396.6	44.2	412	9	AI041451
	7	363.8	40.6	427	10	BF108794
	8	318.6	35.5	488	4	BX529505
c	9	255	28.4	949	12	BI410828
	10	234.4	26.1	333	10	BE983573
	11	224	25.0	297	9	AA772412
	12	210.2	23.4	795	12	BI651936
	13	195.2	21.8	259	10	BE648780
	14	195.2	21.8	327	9	AA968077
c	15	182.6	20.4	362	13	BX089049
	16	169.8	18.9	529	9	AW476657
c	17	154.8	17.3	539	9	AL918370
	18	142	15.8	657	13	BQ078813
	19	111	12.4	256	9	AW762061
	20	107.6	12.0	458	9	AI152190
	21	105.2	11.7	750	29	BZ847665
	22	101	11.3	493	28	BH057870
c	23	100.4	11.2	481	28	AZ987593
	24	90.2	10.1	243	10	BB570162
c	25	85	9.5	512	9	AI073386
	26	81.2	9.1	167	9	AI836531
	27	76.6	8.5	477	10	BE984041
	28	67.2	7.5	765	12	BI522417
	29	67	7.5	769	12	BI413085
	30	66	7.4	751	29	CNS04J6G
	31	64.8	7.2	321	10	BE983721
	32	64.4	7.2	538	9	AL925790
c	33	61	6.8	491	9	AL909688
c	34	60.2	6.7	322	9	AL909689
	35	58.8	6.6	413	14	N62228
	36	58.4	6.5	685	14	CA351220
	37	53	5.9	1630	11	AK051824
	38	51.6	5.8	925	29	CNS0091P
c	39	51.6	5.8	982	13	BX415111
	40	50.8	5.7	647	12	BI960178
	41	50.4	5.6	251	9	AW045376
	42	49.8	5.6	473	12	BI666105
	43	49.8	5.6	658	14	CB059196
	44	49.8	5.6	705	12	BI662853
	45	49.8	5.6	728	13	BQ180353

ALIGNMENTS

RESULT 1
 BI918620 805 bp mRNA linear EST 16-OCT-2001
 LOCUS
 DEFINITION 603176570F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5240969 5',
 mRNA sequence.
 ACCESSION BI918620
 VERSION BI918620.1 GI:16182295
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 805)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11607 row: k column: 18
 High quality sequence start: 2
 High quality sequence stop: 778.
 FEATURES
 source Location/Qualifiers
 1. .805
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5240969"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_121"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
 fetal brains, female age 20 weeks, female age 24 weeks,
 and male age 26 weeks. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 0.7-3.5 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 017. Note:
 this is a NIH MGC Library."
 BASE COUNT 169 a 243 c 263 g 130 t
 ORIGIN
 Query Match 75.1%; Score 674; DB 12; Length 805;
 Best Local Similarity 98.7%; Pred. No. 4.3e-151;
 Matches 732; Conservative 0; Mismatches 5; Indels 5; Gaps 5;

QY 1 ATGAGGCGCGACCCGGCCCCCGGC-TTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTG 59
 |||||
 Db 64 ATGAGGCGCGACCCGGCCCCCGGCGTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTG 123
 QY 60 CTACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGA 119
 |||||
 Db 124 CTACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGA 183
 QY 120 GGGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCC 179
 |||||
 Db 184 GGGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCC 243
 QY 180 GCGCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGG 239
 |||||
 Db 244 GCGCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGG 303
 QY 240 GGGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCA 299
 |||||
 Db 304 GGGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCA 363
 QY 300 GCGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTG- 358
 |||||
 Db 364 GCGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGT 423
 QY 359 CCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTG 418
 |||||
 Db 424 CCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTG 483
 QY 419 ACTGCGCCACCCGGGCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGA 478
 |||||
 Db 484 ACTGCGCCACCCGGGCCAAGTTGAAGAAGATGACGAGCCAGACGGGACAGGTGGGTGAGA 543
 QY 479 AGCAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCA 538
 |||||
 Db 544 AGCAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCA 603
 QY 539 AGGATGGCAAGGAGCTCAACCG-CAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGA 597
 |||||
 Db 604 AGGATGGCAAGGAGCTCAACCGTCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGA 663
 QY 598 AAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGA-GTATGTCTG 656
 |||||
 Db 664 AAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGAGGTATGTCTG 723
 QY 657 CGAGG-CCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCG 715
 |||||
 Db 724 CGAGGCCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGTTTTACGTCAACAGGT 783
 QY 716 TGAGCACCACCCTGTCATCCTG 737
 |||||
 Db 784 TGAGCACCAACCTGTCATCCTG 805

RESULT 2
 BM914622
 LOCUS

BM914622

1047 bp

mRNA

linear

EST 12-MAR-2002

DEFINITION AGENCOURT_6615334 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:5480308
5', mRNA sequence.

ACCESSION BM914622

VERSION BM914622.1 GI:19365001

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1047)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Mark Watson

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM2002 row: p column: 05

High quality sequence stop: 541.

FEATURES

source

Location/Qualifiers

1..1047

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5480308"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_113"

/note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT

263 a 347 c 254 g 183 t

ORIGIN

Query Match 63.1%; Score 565.6; DB 12; Length 1047;

Best Local Similarity 96.8%; Pred. No. 4.6e-125;

Matches 577; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 272 GCTCCTGTGTGCCGCTCGAAAGGAACCAGCGCTACATCTTTTTCCTGGAGCCCACGGAAC 331
|||||

Db 1 GCTCCTGTGTGCCGCTCGAAAGGAACCAGCGCTACATCTTTTTCCTGGAGCCCACGGAAC 60

Qy 332 AGCCCTTAGTCTTTAAGACGGCCTTTGCCCCCTCGATACCAACGGCAAAAATCTCAAGA 391
|||||

Db 61 AGCCCTTAGTCTTTAAGACGGCCTTTGCCCCCTCGATACCAACGGCAAAAATCTCAAGA 120

Qy 392 AAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGA 451
|||||

Db	121	AAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGA	180
Qy	452	AGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTA	511
Db	181	AGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTA	240
Qy	512	ATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACA	571
Db	241	ATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACA	300
Qy	572	TTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGG	631
Db	301	TTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGG	360
Qy	632	TGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAAGGACACCGTCC	691
Db	361	TGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAAGGACACCGTCC	420
Qy	692	GGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCATCCTGGTCGGGGCAGCCCC	751
Db	421	GGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCATCCTGGTCGGGGCAGCCCC	480
Qy	752	GGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCG	811
Db	481	GGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGCCTGCTACTACATCG	540
Qy	812	AGGGCATCAACCAGCTCTCCTGCAAGTGTCCGTGGGATACACCGGGGACAGGTGT	867
Db	541	AGGCCATCAATCAGCTTTCCTGCAAATGTCCCAATGGATTCTTCCGACCAACATGT	596

RESULT 3

BI412864/c

BI412864	LOCUS	BI412864	1041 bp	mRNA	linear	EST	14-AUG-2001
----------	-------	----------	---------	------	--------	-----	-------------

DEFINITION 602988202F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5144016 5', mRNA sequence.

ACCESSION BI412864

VERSION BI412864.1 GI:15173787

KEYWORDS EST.

NETWORKS	5517
SOURCE	Mus musculus (house mouse)

ORGANISM	Mus musculus
----------	--------------

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1041)

NIH-MGC <http://mgc.nci.nih.gov/>.

AUTHORS NIH MSC (http://mgc.nhlbi.nih.gov)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cqapbs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

FEATURES

source

1. .1041

```
/organism="Mus musculus"
```

```
/mol type="mRNA"
```

```
/strain="CZECH II"
```

```
/db xref="taxon:10090"
```

```
/clone="IMAGE:5144016"
```

```
/tissue type="pooled lung tumors"
```

```
/lab host="DH10B (phage-resistant)"
```

```
/clone lib="NCI CGAP Lu33"
```

```

/clone_11b= NotI_SalI_2100
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st

```

strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGGAGCGGCCCTCTGTTTTTTTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors

Double-stranded cDNA was ligated to Eco RI adapter (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library

I and Eco RI sites of the modified pT7-3 vector went through one round of normalization, and was

went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT

247 a 306 c 295 g 193 t

ORIGIN

Query Match 52.1%; Score 467.2; DB 12; Length 1041;

Query Match	32.1%;	Score	10.72;
Best Local Similarity	86.6%;	Pred. No.	1.8e-101;

Best Local Similarity 86.6%; Pred. NO. 1.0e-101;
Matches 563; Conservative 0; Mismatches 78; Indels 9; Gaps 4;

QY 165 CAGCACCCGAGAGCCGCCCGCCTCGGGTCGGGT----GGCGTTGGTAAAGGTGCTGGACA 220
||| ||||||||||||| ||| ||| ||||||||||| 500

Db 656 CACCTCGAGATGCCCGCCCTCGGGTTTCGGTTGGCGTCTTGGTGAAAGGTGCTGGACA 597

Qy 221 AGTGGCCG--CTCCGGAGCGGGGGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTG 278

Dp 596 AGTTGCCGGCTCCCGGATCGGGGGGCTGCAGCGCAGGTGATCAGCGTGGGCTCCTG 537

279 TGTGCCGCTCGAAAGGAACCAGCGCTACATCTT-TTTCCTGGAGCCCACGGAACAGCCCT 337

Qy 279 TGTGCCGCTCGAAAGGAACCAGCGCTACATCTT-TCCTGGAGCCCACCGAGCAGCCCT 307
|| ||||||||||||||||||||||||||||||| || |||||
Db 536 TGCGCCGCTCGAAAGGAACCAGCGCTACATCTTGTTTCCTGGAGCCCACCGAGCAGCCCT 477

Db 536 TGCGCCGCTCGAAAGGAACCAGCGCTACATCTTGTTTCCTGGAGCCACCAGAGAGCCCC 477

QY 338 TAGTCTTTAAGACGGCCTTTGCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGG 397
||||| ||||| |||| | ||||| | |||||
339 TACTTTTTTAAGACGGCCTTTGCCCCGGTCGACCCTACGGCAAATACATCAAGAAAGAGG 417

Db 476 TAGTTTTTAAAGACAGCCTTTTGTCCCGGTCGACCCTACGGCAAATACATCAAGAAAGAGG 417

Qy 398 TGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCC 457
|||||

Db 416 TGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGCTGAAGAAGATGAAGAGCC 35

Qy 458 AGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCC 517
||||| ||| |||||||||||||||| ||||| ||||||||||||| || ||| ||| |||

Db 356 AGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCAAGTGTGAGGCAGCGCGGGAAACCCCC 29

Qy 518 AGCCTTCCTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCA 57

QY 318 AGCCCTCCTATCGCTGGTTCAAGGATGGCAAGGAAGTCAACCGGAGTCGTGATATTGCA 23

Db 296 AGCCCTCCTATCGCTGGTTCAAGGATGGCAAGGAAGTCAACCGGAGTCGTGATATTGCA 23

Qy 578 TCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGG 637
 |||| ||||||| ||||||||||||||| ||||||||||| |||| |||||||
 Db 236 TCAAGTATGGCAATGGCAGAAAGAACTCACGGCTACAGTTCAACAAAGTGAGGGTGGAGG 177
 Qy 638 ACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCC 697
 | || ||||||| |||| ||||||||||||||| ||||||||||| |||||||
 Db 176 ATGCCGGGGAGTACGTCTGTGAGGCCGAGAACATCCTTGGGAAGGACACCGTGAGGGGCC 117
 Qy 698 GGCTTTACGTCAACAGCGTGAGCACCACCCTGTTCATCCTGGTTCGGGGCACGCCCCGGAAGT 757
 | || | ||||||||||||||||||| ||||||||||| || |||||||
 Db 116 GACTCCATGTCAACAGCGTGAGCACCCTGTTCATCCTGGTTCGGGACATGCCCGGAAGT 57
 Qy 758 GCAACGAGACAGCCAAGTCCTA--TTGCGTCAATGGAGGCGTCTGCTACT 805
 |||| ||||| ||||||||||| || || ||||||||||| |||||||
 Db 56 GCAATGAGACCGCCAAGTCCTACCATGTGTGAATGGAGGCGTGTGCTACT 7

RESULT 4

BX281777

LOCUS BX281777 524 bp mRNA linear EST 04-MAR-2003

DEFINITION BX281777 NIH_MGC_121 Homo sapiens cDNA clone IMAGp998K1811607 ;
 IMAGE:5240969, mRNA sequence.

ACCESSION BX281777

VERSION BX281777.1 GI:28612804

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 524)

AUTHORS Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
 Radelof,U., Schneider,D. and Korn,B.

TITLE Human UnigeneSet - RZPD3

JOURNAL Unpublished

COMMENT Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

RZPD; IMAGp998K1811607.

RZPDLIB; I.M.A.G.E. cDNA Clone Collection;

Human UnigeneSet - RZPD3 (RZPDLIB No.972)

[http://www.rzpd.de/CloneCards/cgi-](http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972)

[bin/showLib.pl.cgi/response?libNo=972](http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972) Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany

Tel: +49 30 32639 101

Fax: +49 30 32639 111

www.rzpd.de

This clone is available royalty-free from RZPD;

contact RZPD (clone@rzpd.de) for further information. Seq primer:

M13u, Primer sequence: CGTTGTAAAACGACGGCCAGT.

FEATURES Location/Qualifiers

source 1. 524

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGp998K1811607 ; IMAGE:5240969"

```

/lab_host="DH10B"
/clone_lib="NIH_MGC_121"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH MGC Library."

```

```

BASE COUNT      99 a      174 c      172 g      79 t
ORIGIN

```

```

Query Match          52.1%; Score 467; DB 13; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.6e-101;
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 60
      |||
Db      58 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 117

Qy      61 TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
      |||
Db      118 TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 177

Qy      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
      |||
Db      178 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 237

Qy      181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGCCGCTCCGGAGCGGG 240
      |||
Db      238 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGCCGCTCCGGAGCGGG 297

Qy      241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
      |||
Db      298 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 357

Qy      301 CGCTACATCTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
      |||
Db      358 CGCTACATCTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 417

Qy      361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
      |||
Db      418 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 477

Qy      421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACA 467
      |||
Db      478 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACA 524

```

RESULT 5

AA706226

```

LOCUS      AA706226      549 bp      mRNA      linear      EST 12-JAN-1999
DEFINITION ah28a07.s1 Soares_parathyroid tumor NbHPA Homo sapiens cDNA clone
            1240116 3' similar to TR:P43328 P43328 NEU DIFFERENTIATION FACTOR

```

NDF04 ;, mRNA sequence.

ACCESSION AA706226

VERSION AA706226.1 GI:2716144

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 549)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: similarity on wrong strand
Possible reversed clone: polyT not found
Insert Length: 689 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 451.

FEATURES

source Location/Qualifiers

1. .549
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="1240116"
/tissue_type="parathyroid tumor"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_parathyroid_tumor_NbHPA"
/note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
) with a modified polylinker; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer
[5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCACCAATTTTTTTTTTTTTTTTTTTT
TTTTT-3'], double-stranded cDNA was size selected, ligated
to Eco RI adapters (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of a modified pT7T3
vector (Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
adenomas was kindly provided by Dr. Stephen Marx, National
Institute of Diabetes and Digestive and Kidney Diseases,
NIH."

BASE COUNT 137 a 163 c 156 g 92 t 1 others

ORIGIN

Query Match 45.5%; Score 408.2; DB 9; Length 549;
Best Local Similarity 91.7%; Pred. No. 2.1e-87;

Matches 431; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```

Qy      424 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15  GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 74

Qy      484 TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75  TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 134

Qy      544 GGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAAC 603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      135 GGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAAC 194

Qy      604 TCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCC 663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      195 TCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCC 254

Qy      664 GAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACC 723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      255 GAGAACATCCTGGGGAAGGACACCGTCCGGAGGCCGGCTTTACGTCAACAGCGTGAGCACC 314

Qy      724 ACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGC 783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      315 ACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGNGACAGCCAAGTCCTATTGC 374

Qy      784 GTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTCT 843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      375 GTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGGCACCT 434

Qy      844 GTGGGATACACCGGGGACAGGTGTGAGCAGTTGCAATGGTCAACTTCTC 893
          ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db      435 GGGCTGCACTGCTTAGAACTTGGTACCCAGAGCCACCACTTCCCCATCTC 484

```

RESULT 6

AI041451

LOCUS AI041451 412 bp mRNA linear EST 28-AUG-1998

DEFINITION ow36c02.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
IMAGE:1648898 3' similar to TR:014511 014511 NTAK. ;, mRNA
sequence.

ACCESSION AI041451

VERSION AI041451.1 GI:3280645

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 412)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo,
Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
 Insert Length: 671 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 1.

FEATURES
 source Location/Qualifiers
 1. .412
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1648898"
 /tissue_type="parathyroid tumor"
 /dev_stage="adult"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares_parathyroid_tumor_NbHPA"
 /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
) with a modified polylinker; Site_1: Not I; Site_2: Eco
 RI; 1st strand cDNA was primed with a Not I - oligo(dT)
 primer
 [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCACCAATTTTTTTTTTTTTTTTTTTT
 TTTT-3'], double-stranded cDNA was size selected, ligated
 to Eco RI adapters (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of a modified pT7T3
 vector (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
 adenomas was kindly provided by Dr. Stephen Marx, National
 Institute of Diabetes and Digestive and Kidney Diseases,
 NIH."

BASE COUNT 112 a 108 c 126 g 65 t 1 others
 ORIGIN

Query Match 44.2%; Score 396.6; DB 9; Length 412;
 Best Local Similarity 97.6%; Pred. No. 1.2e-84;
 Matches 402; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy	426	CACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATC	485
Db	1	CACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATC	60
Qy	486	GCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGG	545
Db	61	GCTGAAGTGTGAGGCAGCAGCGATAAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGG	120
Qy	546	CAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAACTC	605
Db	121	CAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAACTC	180
Qy	606	ACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGA	665
Db	181	ACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGA	240

758280-760583, 772104-774407 Soares NbHPA pool 1:
 304776-306311, 320136-322823, 326280-326663 Soares NbHOT
 pool 1: 723720-726407, 739080-740999 Subtraction by Bento
 Soares and M. Fatima Bonaldo."

BASE COUNT 114 a 112 c 145 g 56 t
 ORIGIN

Query Match 40.6%; Score 363.8; DB 10; Length 427;
 Best Local Similarity 91.3%; Pred. No. 8.5e-77;
 Matches 386; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

Qy 363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
    || || | || | | | | | | | | | | | | | | | |
Db 5 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 64

Qy 423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
    |||||
Db 65 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 124

Qy 483 ATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
    |||||
Db 125 ATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 184

Qy 543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602
    |||||
Db 185 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 244

Qy 603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
    |||||
Db 245 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 304

Qy 663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
    |||||
Db 305 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 364

Qy 723 CACCCTGTCATCCTGGTCCGGGGCAGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
    |||||
Db 365 CACCCTGTCATCCTGGTCCGGGGCAGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 424

Qy 783 CGT 785
    |||
Db 425 CGT 427
  
```

RESULT 8
 BX529505

ID BX529505 standard; RNA; EST; 488 BP.

XX

AC BX529505;

XX

SV BX529505.1

XX

DT 27-MAY-2003 (Rel. 75, Created)

DT 27-MAY-2003 (Rel. 75, Last updated, Version 1)

XX

DE RZPD Mus musculus cDNA clone IMAGp998N017639 = IMAGE:3153984 5' EST.

XX

KW EST; expressed sequence tag.
XX
OS Mus musculus (house mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
XX

RN [1]
RP 1-488
RA Heil O., Ebert L., Neubert P., Peters M., Radelof U., Schneider D.,
RA Korn B.;
RT ;
RL Submitted (28-MAY-2003) to the EMBL/GenBank/DDBJ databases.
RL RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Im Neuenheimer
RL Feld 580, D-69120 Heidelberg, Germany
XX

CC RZPD; IMAGp998N017639.
CC RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
CC Mouse UnigeneSet - RZPD2 (RZPDLIB No.981)
CC <http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=981>
CC Contact: Ina Rolfs
CC RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
CC Heubnerweg 6, D-14059 Berlin, Germany
CC Tel: +49 30 32639 101
CC Fax: +49 30 32639 111
CC www.rzpd.de
CC This clone is available royalty-free from RZPD;
CC contact RZPD (clone@rzpd.de) for further information.
CC Seq primer: SP6, Primer sequence: ATTTAGGTGACACTATAG
XX

Key	Location/Qualifiers
source	1. .488
	/db_xref="taxon:10090"
	/note="Cloned unidirectionally. Primer: Oligo dT. Average
	insert 2 kb. Library constructed by Life Technologies,
	catalog #12017-018. Investigators providing samples: Lothar
	Hennighausen/Chu-Xia Deng, NIH Reference for transgenic
	model: Xu et al., Nature Genetics 22, 37-43 (1999). Note:
	this is a NCI_CGAP Library
	< http://www.ncbi.nlm.nih.gov/ncicgap/ >."
	/organism="Mus musculus"
	/clone="IMAGp998N017639"
	/clone_lib="NCI_CGAP_Mam3 mammary tumor"
	/lab_host="DH10B"

SQ Sequence 488 BP; 129 A; 116 C; 149 G; 94 T; 0 other;
Query Match 35.5%; Score 318.6; DB 4; Length 488;
Best Local Similarity 86.7%; Pred. No. 6.1e-66;
Matches 351; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY	469	GTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTAC	528
Db	1	GTGGGTGAGAAGCAGTCGCTCAAGTGTGAGGCAGCGCGGAAACCCAGCCCTCCTAT	60
QY	529	CGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGC	588

Db	61	CGCTGGTTC AAGGATGGCAAGGAACTCAACCGGAGTCGTGATATTCGCATCAAGTATGGC	120
QY	589	AACGGCAGAAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAG	648
Db	121	AATGGCAGAAAAGAACTCACGGCTACAGTTCAACAAAGTGAGGGTGGAGGATGCCGGGGAG	180
QY	649	TATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTC	708
Db	181	TACGTCTGTGAGGCCGAGAACATCCTTGGAAGGACACCGTGAGGGGCCGACTCCATGTC	240
QY	709	AACAGCGTGAGCACCACCCTGTTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACA	768
Db	241	AACAGCGTGAGCACCCTGTTCATCCTGGTTCGGGACATGCCCGGAAGTGCAATGAGACC	300
QY	769	GCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTC	828
Db	301	GCCAAGTCCTACTGTGTGAATGGAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTC	360
QY	829	TCCTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGTTCAGCAG	873
Db	361	TCCTGCAAAATGTCCAAACGGATTCTTCGGACAGAGATGTTTGGAG	405

RESULT 9

BI410828/c

LOCUS

BI410828

949 bp

mRNA

linear

EST 14-AUG-2001

DEFINITION

00298573411 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5119065 5',
mRNA sequence.

ACCESSION

BI410828

VERSION

BI410828.1 GI:15171751

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 949)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM11290 row: d column: 10

High quality sequence start: 28

High quality sequence stop: 919.

FEATURES

Location/Qualifiers

source

1. .949

```
/organism="Mus musculus"
```

```
/mol type="mRNA"
```

```
/strain="CZECH TT"
```

/db_xref="taxon:10090"
 /clone="IMAGE:5119065"
 /tissue_type="pooled lung tumors"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NCI_CGAP_Lu33"
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5' TGTACCAATCTGAAGTGGGAGCGGCCGCTCTGTTTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 171 a 261 c 269 g 248 t
 ORIGIN

Query Match 28.4%; Score 255; DB 12; Length 949;
 Best Local Similarity 80.0%; Pred. No. 1.3e-50;
 Matches 388; Conservative 0; Mismatches 85; Indels 12; Gaps 7;

QY	397	GTGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGC	456
Db	947	GTGGGCCAGATCCTGGGCACTG-CTGCGCCACCCGCCCAAA-CTGAAGAAGATGAAGA-C	891
QY	457	CAGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCC	516
Db	890	CAAACCAGAAGAGTCGGTGAGAACAGTTCGCTCAAGTGTGAGGCACGGCCGGGGAAACCC	831
QY	517	CAGCCTTCCTACC-----GTTGGTTCAAGGATGGC-AAGGAGCTCAACCGCAGCCGAGAC	570
Db	830	CCCCACCCCTCCCTATCGCTGGTTTCAAGGATGGCAAAGGAAGTCAACCGGAGTCGTGAT	771
QY	571	ATTTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAG	630
Db	770	ATTTCGCATCAAGTATGCCAATGGCAGAAAGAACTCACGGCTACAGTTCAACAAAAGTGAG	711
QY	631	GT--GGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCG	688
Db	710	GTTGGAGGATTGCCGGGGAGTACGTCTGTGAGGCCGAGAACATCCTTGGGAAGGACACCG	651
QY	689	TCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTTCATCCTGGTCGGGGCAGC	748
Db	650	T-GAGGGCCGACTCCATGTCAACAGCGTGAGCACCCTGTTCATCCTGGTCGGGGACATG	592
QY	749	CCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACA	808
Db	591	CCCGGAAGTGCAATGAGACCGCCAAGTCCTACTGTGTGAATGGAGGCGTGTGCTACTACA	532
QY	809	TCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGTC	868
Db	531	TCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAACGGATTCTTCGGACAGAGATGTT	472
QY	869	AGCAG	873
Db	471	TGGAG	467

RESULT 10
 BE983573
 LOCUS BE983573 333 bp mRNA linear EST 29-APR-2002
 DEFINITION UI-M-CG0p-bgi-c-07-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
 UI-M-CG0p-bgi-c-07-0-UI 3', mRNA sequence.
 ACCESSION BE983573
 VERSION BE983573.1 GI:10654893
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 333)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mEST@mail.nih.gov
 Oligo-dT track not found, Not I site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
 clones from RESEARCH GENETICS. It should be noted that Bento Soares
 is generating a small number of additional specialized
 non-redundant arrays of BMAP cDNAs whose availability will be
 considered under appropriate and limited collaborative arrangements
 The tissue for this library was contributed by Dr. Xin-Yuan Fu,
 Yale University School of Medicine The following repetitive
 elements were found in this cDNA sequence: 15-105,
 >GC_rich#Low_complexity
 Seq primer: M13 Forward
 POLYA=No.
 FEATURES
 source Location/Qualifiers
 1. .333
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-CG0p-bgi-c-07-0-UI"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH_BMAP_Ret4_S2"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The
 NIH_BMAP_Ret4_S2 library is a subtracted library,
 ultimately derived from mouse retina tissue libraries at
 various stages of development. For a detailed description
 of the library from which this clone was derived, please
 visit our web site at brainest.eng.uiowa.edu. The tissue

for this library was contributed by Dr. Xin-Yuan Fu, Yale
University School of Medicine
TAG_SEQ=None found"

BASE COUNT 47 a 124 c 122 g 40 t
ORIGIN

Query Match 26.1%; Score 234.4; DB 10; Length 333;
Best Local Similarity 95.6%; Pred. No. 8.4e-46;
Matches 241; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
QY      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCTCGCTCGCCTGC 60
      |||
Db      82 ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 141
      |||
QY      61 TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
      |||
Db     142 TACTCGCCAGCCTCAAGTCGGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 201
      |||
QY     121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
      |||
Db     202 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCG 261
      |||
QY     181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
      |||
Db     262 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 321
      |||
QY     241 GGGCTGCAGCGC 252
      |||
Db     322 GGGCTGCAGCGC 333
```

RESULT 11

AA772412

LOCUS

AA772412

297 bp

mRNA

linear

EST 31-DEC-1998

DEFINITION

ai44e12.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
1359886 3' similar to TR:P43328 P43328 NEU DIFFERENTIATION FACTOR
NDF04 ;, mRNA sequence.

ACCESSION

AA772412

VERSION

AA772412.1 GI:2824195

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 297)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL

Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

Possible reversed clone: similarity on wrong strand
Possible reversed clone: polyT not found
Insert Length: 667 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 267.

source

1. .297

/mol_type="mRNA"

```
/clone="1359886"
```

```
/dev_stage="adult"
```

```
/clone_lib="Soares_parathyroid_tumor NbHPA"
```

```

/note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
) with a modified polylinker; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer

```

[5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCACCAATTTTTTTTTTTTTTTTTTTT
TTTTT-3'], double-stranded cDNA was size selected, ligated
to Eco RI adapters (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of a modified pT7T3
vector (Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M. Fatima Bonaldo. RNA from sporadic parathyroid
adenomas was kindly provided by Dr. Stephen Marx, National
Institute of Diabetes and Digestive and Kidney Diseases,
NIH."

87 a 68 c 105 g 37 t

25.0%; Score 224; DB 9; Length 297;

Best Local Similarity 92.2%; Pred. No. 2.5e-43;
Matches 236;

Matches 236; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Db
42 GAGCCCGATCCC GGGAGAAAGCCACCCGGCCAAGTTGAAGAAGATGAAGAGCCAGACGGG 101

Db
102 ACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCGGTGAATCCCCAGCCTTC 161

Db 162 CTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATA 221

Db 222 TGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGG 281

Db 282 GGAGTATGTCTGCGAG 297

Db 282 GGAGTATGTCTGCGAG 297

RESULT 12

BI651936

LOCUS

BI651936

795 bp

mRNA

linear

EST 12-SEP-2001

DEFINITION

603298677F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5339251 5', mRNA sequence.

ACCESSION

BI651936

VERSION

BI651936.1 GI:15566172

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 795)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11861 row: j column: 20

High quality sequence stop: 795.

FEATURES

source

Location/Qualifiers

1. .795

/organism="Mus musculus"

/mol_type="mRNA"

/strain="129,C57BL/6J,FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:5339251"

/tissue_type="tumor, gross tissue"

/dev_stage="10 months"

/lab_host="DH10B"

/clone_lib="NCI_CGAP_Mam3"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigators

providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH

Reference for transgenic model: Xu et al., Nature Genetics

22, 37-43 (1999)."

BASE COUNT

204 a 226 c 219 g 146 t

ORIGIN

Query Match

23.4%; Score 210.2; DB 12; Length 795;

Best Local Similarity 86.2%; Pred. No. 7e-40;

Matches 244; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY

592 GGCAGAAAGAACTCACGAC-TACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTA 650
|||||

Db

1 GGCAGAAAGAACTCACGGCTTACAGTTCAACAAGGTGAGGGTGGAGGATGCCGGGGAGTA 60

QY 651 TGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAA 710
 Db 61 CGTCTGTGAGGCCGAGAACATCCTTGGGAAGGACACCGTGAGGGGCCGACTCCATGTCAA 120
 QY 711 CAGCGTGAGCACCACCCTGTCTATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGC 770
 Db 121 CAGCGTGAGCACCCTGTCTATCCTGGTCGGGACATGCCCCGGAAGTGCAATGAGACCGC 180
 QY 771 CAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTC 830
 Db 181 CAAGTCCTACTGTGTGAATGGAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTC 240
 QY 831 CTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873
 Db 241 CTGCAATGTCCAAACGGATTCTTCGGACAGAGATGTTTGGAG 283

RESULT 13

BE648780

LOCUS

BE648780

259 bp

mRNA

linear

EST 06-SEP-2000

DEFINITION

UI-M-BH2.2-aop-b-12-0-UI.r1 NIH_BMAP_M_S3.2 Mus musculus cDNA clone
 UI-M-BH2.2-aop-b-12-0-UI 5', mRNA sequence.

ACCESSION

BE648780

VERSION

BE648780.1 GI:9974601

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 259)

REFERENCE

AUTHORS

Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE

Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL

Genome Res. 6 (9), 791-806 (1996)

MEDLINE

97044477

PUBMED

8889548

COMMENT

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mEST@mail.nih.gov

cDNA Library Preparation: M.B. Soares Lab Clone distribution:
 Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
 should be noted that Bento Soares is generating a small number of
 additional specialized non-redundant arrays of BMAP cDNAs whose
 availability will be considered under appropriate and limited
 collaborative arrangements

Seq primer: M13 Reverse.

FEATURES

source

Location/Qualifiers

1. .259

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-BH2.2-aop-b-12-0-UI"

Qy	701	TTTACGTCAACAGCGTGAGCACCACCTGT	CATCCTGGTCGGGGGCACGCCCGGAAGTGCA	760
Db	61	TCCATGTCAACAGCGTGAG-ACCACTCTGT	CATCCTGGTCGGGACATGCCCGGAAGTGCA	119
Qy	761	ACGAGACAGCCAAGTCTATTGCGTCAATGG	AGGCGTCTGCTACTACATCGAGGGCATCA	820
Db	120	ATGAGACCGCCAAGTCTACTGTGTGAATGG	AGGCGTGTGCTACTACATCGAGGGCATCA	179
Qy	821	ACCAGCTCTCCTGCAAGTGTCTGTGGGATA	CACCGGGGACAGGTGTGAGCAGTTCGCAA	880
Db	180	ACCAGCTCTCCTGCAAGTGTCTGTGGGATA	CACCGGGGACAGGTGTGAGCAGTTCGCAA	239
Qy	881	TGGTCAACTTCTCCTA		896
Db	240	TGGTCAACATCTCCAA		255

```

RESULT 15
BX089049/c
LOCUS BX089049 362 bp mRNA linear EST 23-JAN-2003
DEFINITION BX089049 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
IMAGp998M133119 ; IMAGE:1240116, mRNA sequence.
ACCESSION BX089049
VERSION BX089049.1 GI:27825909
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 362)
AUTHORS Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
Radelof,U., Schneider,D. and Korn,B.
TITLE Human UnigeneSet - RZPD3
JOURNAL Unpublished
COMMENT Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGp998M133119.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human UnigeneSet - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES
source Location/Qualifiers
1. .362
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998M133119 ; IMAGE:1240116"
/tissue_type="parathyroid tumor"

```


OM nucleic - nucleic search, using sw model

Run on: January 14, 2004, 07:12:21 ; Search time 3588 Seconds
(without alignments)
10227.407 Million cell updates/sec

Title: US-09-864-675-3

Perfect score: 897

Sequence: 1 atgaggcgcgacccggcccc.....caatgggtcaacttctcctaa 897

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*

28: em_un:*
 29: em_vi:*
 30: em_htg_hum:*
 31: em_htg_inv:*
 32: em_htg_other:*
 33: em_htg_mus:*
 34: em_htg_pln:*
 35: em_htg_rod:*
 36: em_htg_mam:*
 37: em_htg_vrt:*
 38: em_sy:*
 39: em_htgo_hum:*
 40: em_htgo_mus:*
 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result

Query

No.

Score

Match

Length

DB

ID

Description

1

849

94.6

3020

9

AB005060

AB005060 Homo sapi

2

835.4

93.1

1884

6

AR098145

AR098145 Sequence

3

835.4

93.1

1884

6

AR116617

AR116617 Sequence

4

784

87.4

3441

6

AR072052

AR072052 Sequence

5

738.6

82.3

993

6

AR072053

AR072053 Sequence

6

737

82.2

2947

10

D89995

D89995 Rattus sp.

7

737

82.2

3076

6

E16456

E16456 Rat mRNA fo

8

737

82.2

3077

10

D89996

D89996 Rattus sp.

9

427.8

47.7

1476

6

AR098146

AR098146 Sequence

10

427.8

47.7

1476

6

AR116618

AR116618 Sequence

11

427.8

47.7

2268

6

AR098155

AR098155 Sequence

12

427.8

47.7

2268

6

AR116627

AR116627 Sequence

13

425

47.4

2188

10

AB001576

AB001576 Rattus sp

14

424.8

47.4

118504

9

AC094080

AC094080 Homo sapi

c 15

424.8

47.4

152838

2

AC011589

AC011589 Homo sapi

16

424.8

47.4

170797

9

AC011379

AC011379 Homo sapi

17

424.8

47.4

210675

2

AC026272

AC026272 Homo sapi

18

424

47.3

1054

6

AX406616

AX406616 Sequence

19

424

47.3

1054

9

HS2NRG01

AF119151 Homo sapi

20

405.4

45.2

1607

6

AR098144

AR098144 Sequence

21

405.4

45.2

1607

6

AR116616

AR116616 Sequence

22

405.4

45.2

2467

6

AR098143

AR098143 Sequence

23

405.4

45.2

2467

6

AR116615

AR116615 Sequence

24

387.2

43.2

140307

2

AC131191

AC131191 Mus muscu

25

384

42.8

253462

2

AC096477

AC096477 Rattus no

26

216.2

24.1

1207

6

AR072054

AR072054 Sequence

27

173

19.3

419

6

AX406617

AX406617 Sequence

28

173

19.3

419

9

HS2NRG02

AF119152 Homo sapi

29

173

19.3

120236

9

AC008523

AC008523 Homo sapi

c 30

173

19.3

189050

9

AC008667

AC008667 Homo sapi

31

142.6

15.9

85703

2

AC020830

AC020830 Mus muscu

c 32

142.6

15.9

191101

2

AC127350

AC127350 Mus muscu

33

139.4

15.5

226038

2

AC106592

AC106592 Rattus no

	34	139.4	15.5	273080	2	AC098540	AC098540 Rattus no
c	35	139.4	15.5	302176	2	AC096479	AC096479 Rattus no
	36	130.2	14.5	163	10	AY227026	AY227026 Mus muscu
	37	124.6	13.9	493	6	AX406618	AX406618 Sequence
	38	124.6	13.9	493	9	HS2NRG03	AF119153 Homo sapi
	39	124	13.8	350	6	AX406619	AX406619 Sequence
	40	124	13.8	350	9	HS2NRG04	AF119154 Homo sapi
c	41	109.6	12.2	85703	2	AC020830	AC020830 Mus muscu
	42	108.4	12.1	206683	2	BX323592	BX323592 Danio rer
	43	108.4	12.1	220700	2	BX005008	BX005008 Danio rer
	44	97	10.8	172	10	D89997	D89997 Rattus sp.
	45	84.8	9.5	240	10	AY227025	AY227025 Mus muscu

ALIGNMENTS

RESULT 1
AB005060

LOCUS AB005060 3020 bp mRNA linear PRI 14-NOV-1997

DEFINITION Homo sapiens mRNA for NTAK, complete cds.

ACCESSION AB005060

VERSION AB005060.1 GI:2626738

KEYWORDS NTAK.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)

AUTHORS Higashiyama, S., Horikawa, M., Yamada, K., Ichino, N., Nakano, N., Nakagawa, T., Miyagawa, J., Matsushita, N., Nagatsu, T., Taniguchi, N. and Ishiguro, H.

TITLE A novel brain-derived member of the epidermal growth factor family that interacts with ErbB3 and ErbB4

JOURNAL J. Biochem. 122 (3), 675-680 (1997)

MEDLINE 98006324

PUBMED 9348101

REFERENCE 2 (bases 1 to 3020)

AUTHORS Ishiguro, H.

TITLE Direct Submission

JOURNAL Submitted (24-JUN-1997) Hiroshi Ishiguro, Fujita Health University, ICMS; 1-98, kutsukake-cho, Toyoake, Aichi 470-11, Japan (E-mail:hishi@fujita-hu.ac.jp, Tel:0562-93-9393, Fax:0562-93-8831)

FEATURES

source Location/Qualifiers

1. .3020
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="SK-N-SH"
/cell_type="neuroblastoma"

CDS 226. .2778
/codon_start=1
/product="NTAK"
/protein_id="BAA23417.1"
/db_xref="GI:2626739"
/translation="MRQVCCSALPPPPLEKGRCSSYSDSSSSSSERSSSSSSSSSESG
SSSRSSSNSSISRPAAPPEPRPQQQPQPRSPAARRAAARSRAAAAGGMRRDPAPGFS"

MLLFGVSLACYSPSLKSVQDQAYKAPVVVEGKVQGLVPAGGSSSNSTREPPASGRVAL
VKVLDKWPLRSGGLQREQVISVGSCVPLERNQRYIFFLEPTEQPLVFKTAFAPLDTNG
KNLKKEVGKILCTDCATRPKLKKMKSQTGQVGEKQSLKCEAAAGNPQPSYRWFKDGKE
LNRSRDIRIKYGNRKNRSRLQFNKVVEDAGEYVCEAENILGKDTVGRGLVNSVSTT
LSSWSGHARKCNETAKSYCVNGGVCYYIEGINQLSCKCPNGFFGQRCLEKLPLRLYMP
DPKQKAEELYQKRVLTITGICVALLVVGIVCVVAYCKTKKQKQMHNLHRLQNMCPAHQ
NRSLANGPSPHRLDPEEIOMADYISKVNPATDHVIRRETETTFSGSHSCSPSHHCSTA
TPTSSHRHESHTWSLSESLTSDSQSGIMLSSVGTSKCNSPACVEARARRAAAYNLE
ERRRATAPPYHDSVDSLSDSPHSERYVSALTTPARLSPVDFHYSLATQVPTFEITSPN
SAHAVSLPPAAPISYRLAEQQPLLRRHPAPFGPGPGPGADMQRSYDSYYPAA
GPGPRRGTCALGGSGLGSLPASFPRIPEDEYETTQECAPPPPRPRARGASRRTSAGP
RRWRRSRLNGLAAQRARAARDSLSLSSGSGGGSASASDDDDADDGALAAESTPFLGL
RGAHDALRSDSPPLCPAADSRTYYSLDSHSTRASSRHSRGPppPRAKQDSAPL"

polyA_site

3020

/note="39 A nucleotides"

BASE COUNT 615 a 1015 c 937 g 453 t
ORIGIN

Query Match 94.6%; Score 849; DB 9; Length 3020;
Best Local Similarity 98.3%; Pred. No. 7.3e-152;
Matches 858; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTGC	60
Db	502	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTGC	561
QY	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	562	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	621
QY	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	622	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	681
QY	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	682	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	741
QY	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	742	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	801
QY	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	802	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	861
QY	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	862	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	921
QY	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	922	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	981
QY	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	982	CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	1041

QY 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 600
 |||
 Db 1042 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 1101
 QY 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
 |||
 Db 1102 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 1161
 QY 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
 |||
 Db 1162 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 1221
 QY 721 ACCACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
 |||
 Db 1222 ACCACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 1281
 QY 781 TGGCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
 |||
 Db 1282 TGGCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 1341
 QY 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873
 || ||| | || | || || ||
 Db 1342 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 1374

RESULT 2

AR098145
 LOCUS AR098145 1884 bp DNA linear PAT 14-FEB-2001
 DEFINITION Sequence 5 from patent US 6074841.
 ACCESSION AR098145
 VERSION AR098145.1 GI:12807402
 KEYWORDS .
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 1884)
 AUTHORS Gearing,D.P. and Busfield,S.J.
 TITLE Don-1 gene and polypeptides and uses therefor
 JOURNAL Patent: US 6074841-A 5 13-JUN-2000;
 FEATURES Location/Qualifiers
 source 1. .1884
 /organism="unknown"
 BASE COUNT 426 a 607 c 560 g 291 t
 ORIGIN

Query Match 93.1%; Score 835.4; DB 6; Length 1884;
 Best Local Similarity 98.1%; Pred. No. 2.9e-149;
 Matches 856; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTGC 60
 |||
 Db 218 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTGC 277
 QY 61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
 |||
 Db 278 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 337

QY 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
 |||
 Db 338 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 397
 QY 181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
 |||
 Db 398 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 457
 QY 241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
 |||
 Db 458 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 517
 QY 301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
 |||
 Db 518 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 577
 QY 361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
 ||| |||
 Db 578 CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC 636
 QY 421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
 |||
 Db 637 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 696
 QY 481 CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
 |||
 Db 697 CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 756
 QY 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG 600
 |||
 Db 757 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG 816
 QY 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
 |||
 Db 817 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 876
 QY 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
 |||
 Db 877 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 936
 QY 721 ACCACCCTGTCATCCTGGTTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTAT 780
 |||
 Db 937 ACCACCCTGTCATCCTGGTTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTAT 996
 QY 781 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
 |||
 Db 997 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAATGT 1056
 QY 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873
 || ||| ||| ||| |||
 Db 1057 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 1089

RESULT 3
 AR116617
 LOCUS

AR116617

1884 bp

DNA

linear

PAT 16-MAY-2001

DEFINITION Sequence 5 from patent US 6133423.
 ACCESSION AR116617
 VERSION AR116617.1 GI:14096939
 KEYWORDS .
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 1884)
 AUTHORS Gearing,D.P. and Busfield,S.J.
 TITLE Don-1 gene and polypeptides and uses therefor
 JOURNAL Patent: US 6133423-A 5 17-OCT-2000;
 FEATURES Location/Qualifiers
 source 1. .1884
 /organism="unknown"
 BASE COUNT 426 a 607 c 560 g 291 t
 ORIGIN

Query Match 93.1%; Score 835.4; DB 6; Length 1884;
 Best Local Similarity 98.1%; Pred. No. 2.9e-149;
 Matches 856; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC	60
Db	218	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC	277
Qy	61	TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	278	TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	337
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	338	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	397
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	398	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	457
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	458	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	517
Qy	301	CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	518	CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	577
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	578	CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC	636
Qy	421	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	637	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	696
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	697	CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	756

Qy 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 600
 |||
 Db 757 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 816

Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
 |||
 Db 817 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 876

Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
 |||
 Db 877 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 936

Qy 721 ACCACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
 |||
 Db 937 ACCACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 996

Qy 781 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
 |||
 Db 997 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 1056

Qy 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873
 || ||| | || | || || ||
 Db 1057 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 1089

RESULT 4

AR072052

LOCUS AR072052 3441 bp DNA linear PAT 18-FEB-2000

DEFINITION Sequence 1 from patent US 5912326.

ACCESSION AR072052

VERSION AR072052.1 GI:7222940

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 3441)

AUTHORS Chang,H.

TITLE Cerebellum-derived growth factors

JOURNAL Patent: US 5912326-A 1 15-JUN-1999;

FEATURES Location/Qualifiers

source 1. .3441

/organism="unknown"

BASE COUNT 777 a 1057 c 1015 g 592 t

ORIGIN

Query Match 87.4%; Score 784; DB 6; Length 3441;

Best Local Similarity 92.2%; Pred. No. 1.8e-139;

Matches 826; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTGC 60
 |||
 Db 180 ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 239

Qy 61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
 |||
 Db 240 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 299

Qy 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
 |||||
 Db 300 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGTTCCAGCTCTAACAGCACCCGAGAGCCT 359
 Qy 181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
 |||||
 Db 360 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 419
 Qy 241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
 |||||
 Db 420 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 479
 Qy 301 CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
 |||||
 Db 480 CGCTACATCTTTTTCCTGGAGCCCACCGAGCAGCCCTTAGTCTTTAAGACAGCCTTTGCC 539
 Qy 361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
 || |||||
 Db 540 CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 599
 Qy 421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
 |||||
 Db 600 TGCGCAACCCGGCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG 659
 Qy 481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
 || |||||
 Db 660 CAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCCCAGCCCTCCTATCGATGGTTCAAG 719
 Qy 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAAG 600
 || |||||
 Db 720 GACGGCAAGGAGCTCAACCGGAGTCGTGACATTTCGCATCAAGTATGGCAACGGCAGAAAAG 779
 Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
 |||||
 Db 780 AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG 839
 Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
 || |||||
 Db 840 GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC 899
 Qy 721 ACCACCCCTGTCATCCTGGTTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
 |||||
 Db 900 ACCACTCTGTCTGCTGGTTCGGGGCACGCCCGGAAGTGCAATGAGACAGCCAAGTCCTAC 959
 Qy 781 TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
 || |||||
 Db 960 TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAAGTGT 1019
 Qy 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTCCTA 896
 |||||
 Db 1020 CCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTCAA 1075

RESULT 5

AR072053

LOCUS AR072053 993 bp DNA linear PAT 18-FEB-2000
 DEFINITION Sequence 3 from patent US 5912326.

ACCESSION AR072053
 VERSION AR072053.1 GI:7222941
 KEYWORDS .
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 993)
 AUTHORS Chang,H.
 TITLE Cerebellum-derived growth factors
 JOURNAL Patent: US 5912326-A 3 15-JUN-1999;
 FEATURES Location/Qualifiers
 source 1. .993
 /organism="unknown"
 BASE COUNT 230 a 271 c 311 g 181 t
 ORIGIN

Query Match 82.3%; Score 738.6; DB 6; Length 993;
 Best Local Similarity 90.4%; Pred. No. 8.6e-131;
 Matches 789; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC	60
Db	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC	60
Qy	61	TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	61	TACTCGCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	121	GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT	180
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGCCGCTCCGGAGCGGG	240
Db	181	CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGCCGCTCCGGAGCGGG	240
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG	300
Qy	301	CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	301	CGCTACATCTTTTTCCTGGAGCCCACCGAGCAGCCCTTAGTCTTTAAGACAGCCTTTGCC	360
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	361	CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	421	TGCGCAACCCGGCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG	480
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	481	CAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCCCAGCCCTCCTATCGATGGTTCAAG	540
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG	600

CDS

```
/cell_line="PC12"  
/cell_type="pheochromocytoma"  
79. .2685  
/codon_start=1  
/product="NTAK alpha1"  
/protein_id="BAA23344.1"  
/db_xref="GI:2605630"  
/translation="MRQVCCSALPPPLEKARCSSYSYSDSSSSSSSSNNSSSSTSSRSS  
SRSSSRSSRGSTTTTSSSENSGNSGSI FRPAAPPEPRPQPPQPRSPAARRAAARS  
AAAAGGMRRDPAPGSSMLLFGVSLACYSPLKSVQDQAYKAPVVVEGKVQGLAPAGGS  
SSNSTREPPASGRVALVKVLDKWPLRSGGLQREQVISVGSCAPLERNQRYIFFLEPTE  
QPLVFKTAFAPVDPNGKNIKKEVGKILCTDCATRPKLKMKMSQTGEVGEKQSLKCEAA  
AGNPQPSYRWFKD GKELNRSRDIRIKYGNRKN SRLQFNKVVEDAGEYVCEAENILG  
KDTVGRGLHVNVSSTLSSWSGHARKCNETAKSYCVNGGV CYIIEGINQLSCKCPNGF  
FGQRCLEKLPLRLYMPDPKQKHLGFELKEAEELYQKRVLTITGICVALLVVGIVCVVA  
YCKTKKQRRQMHHHLRQNMCPAHQNRSLANGPSHPRLDPEEIOMADYISKNV PATDHV  
IRREAETTFSGSHSCSPSHHCSTATPTSSHRHESHTWSLERSSELTSDSQSGIMLSSV  
GTSKCN SPACVEARARRAAAYSQEERRRAAMPYHDSIDSLRDSPHSERVVSALTTPA  
RLSPVDFHYSLATQVPTFEITSPNSAHAVSLPPAAPISYRLAEQQPLL RHPAPPGP GP  
GPGADMQRSYDSYYYPAAGPGPRRGACALGGS LGS LPA SPFRI PEDDEYETQECAPF  
PPRPRTRGASRRTSAGPRRWRRLNGLAAQRARAARDSLSLSSGSGCGSASASDDD  
ADDADGALAAESTPFLGLRAAHDALRSDSPPLCPAADSRTYISLDSHSTRASSRHSRG  
PPTRAKQDSGPL"
```

BASE COUNT 665 a 945 c 895 g 442 t
ORIGIN

Query Match 82.2%; Score 737; DB 10; Length 2947;
Best Local Similarity 90.3%; Pred. No. 1.6e-130;
Matches 788; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTGC	60
Db	403	ATGAGGCGCGACCCGGCCCCCGGCTCCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC	462
Qy	61	TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	463	TACTCGCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	522
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	523	GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCAGCTCTAACAGCACCCGAGAGCCT	582
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	583	CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	642
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	643	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG	702
Qy	301	CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	703	CGCTACATCTTTTTCCTGGAGCCCACCGAGCAGCCCTTAGTCTTTAAGACAGCCTTTGCC	762
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	763	CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	822

Qy	421	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	823	TGCGCAACCCGGCCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG	882
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCTTACCGTTGGTTCAAG	540
Db	883	CAGTCGCTCAAGTGTGAGGCAGCGCGGGGAACCCCAGCCCTCTATCGATGGTTCAAG	942
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	943	GACGGCAAGGAGCTCAACCGGAGTCGTGACATTCGCATCAAGTATGGCAACGGCAGAAAG	1002
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	1003	AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG	1062
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	1063	GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC	1122
Qy	721	ACCACCTGTTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	1123	ACCACTCTGTCTCCTGGTCGGGGCACGCCCGGAAGTGCAATGAGACAGCCAAGTCCTAC	1182
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT	840
Db	1183	TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAATGT	1242
Qy	841	CCTGTGGGATACACCGGGGACAGGTGTCAGCAG	873
Db	1243	CCAAACGGATTCTTCGGACAGAGATGTTTGGAG	1275

E16456

VERSION E16456.1 GI:5711139

SOURCE Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

AUTHORS Higashiyama, S., Taniguchi, N., Ishiguro, K. and Nagatsu, T.

JOURNAL Patent: JP 1998179166-A 1 07-JUL-1998;

COMMENT OS Rattus sp. (rat)

PD 07-JUL-1998

PI HIGASHIYAMA SHIGEKI, TANIGUCHI NAOYUKI, ISHIGURO KEIJI, PI

NAGATSU TOSHIHARU

PC C12N15/09,C07K14/705,C07K16/28,C12N5/10,C12N15/02,C12P21/02,
 PC C12P21/08,
 PC C12Q1/68,G01N33/53,G01N33/566//A61K31/70,A61K38/46,A61K39/395,
 PC A61K48/00,
 PC C07H21/04,(C12N5/10,C12R1:91),(C12P21/02,C12R1:91); CC
 strandedness: Double;
 CC topology: Linear;
 FH Key Location/Qualifiers
 FH
 FT source 1. .3076
 FT /organism='Rattus sp.'
 FT /cell_line='PC12'
 FT CDS 232. .2814
 FT /product='NTAK protein'.

FEATURES Location/Qualifiers
 source 1. .3076
 /organism="Rattus sp."
 /mol_type="genomic DNA"
 /db_xref="taxon:10118"

BASE COUNT 673 a 996 c 944 g 463 t
 ORIGIN

Query Match 82.2%; Score 737; DB 6; Length 3076;
 Best Local Similarity 90.3%; Pred. No. 1.6e-130;
 Matches 788; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC	60
Db	556	ATGAGGCGCGACCCGGCCCCCGGCTCCTCGATGCTGCTCTTCGGTGTGTCACCTCGCCTGC	615
Qy	61	TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	616	TACTCGCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	675
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	676	GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCAGCTCTAACAGCACCCGAGAGCCT	735
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGCCGCTCCGGAGCGGG	240
Db	736	CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGCCGCTCCGGAGCGGG	795
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	796	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG	855
Qy	301	CGCTACATCTTTTCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	856	CGCTACATCTTTTCTGGAGCCCACCGAGCAGCCCTTAGTCTTTAAGACAGCCTTTGCC	915
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	916	CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	975
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480


```

      |||
Db      857 CGCTACATCTTTTCTGGAGCCCAACGAGCAGCCCTTAGTTTTAAAGACAGCCTTTGCC 916
      |||
QY      361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
      |||
Db      917 CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 976
      |||
QY      421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
      |||
Db      977 TGCGCAACCCGGCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG 1036
      |||
QY      481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
      |||
Db      1037 CAGTCGCTCAAGTGTGAGGCGGCGGGGGAACCCCCAGCCCTCCTATCGATGGTTCAAG 1096
      |||
QY      541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 600
      |||
Db      1097 GACGGCAAGGAGCTCAACCGGAGTCGTGACATTTCGCATCAAGTATGGCAACGGCAGAAAG 1156
      |||
QY      601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
      |||
Db      1157 AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG 1216
      |||
QY      661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
      |||
Db      1217 GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC 1276
      |||
QY      721 ACCACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
      |||
Db      1277 ACCACTCTGTCTCCTGGTTCGGGGCACGCCCCGGAAGTGCAATGAGACAGCCAAGTCCTAC 1336
      |||
QY      781 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
      |||
Db      1337 TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAAATGT 1396
      |||
QY      841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873
      |||
Db      1397 CCAAACGGATTCTTCGGACAGAGATGTTTGAG 1429

```

RESULT 9

AR098146

LOCUS AR098146 1476 bp DNA linear PAT 14-FEB-2001

DEFINITION Sequence 7 from patent US 6074841.

ACCESSION AR098146

VERSION AR098146.1 GI:12807403

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1476)

AUTHORS Gearing, D.P. and Busfield, S.J.

TITLE Don-1 gene and polypeptides and uses therefor

JOURNAL Patent: US 6074841-A 7 13-JUN-2000;

FEATURES Location/Qualifiers

source 1. .1476

/organism="unknown"

BASE COUNT 335 a 473 c 452 g 216 t
ORIGIN

Query Match 47.7%; Score 427.8; DB 6; Length 1476;
Best Local Similarity 89.8%; Pred. No. 1.4e-71;
Matches 459; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```
Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
          || || |  || |  |  |  |  |||  || |  || |  |  ||
Db      98  CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
          |||||
Db     158  AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
          |||||
Db     218  ATCGCTGAAGTGTGAGGCAGCAGCCGGAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAGAA 602
          |||||
Db     278  TGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAGAA 337

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
          |||||
Db     338  CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
          |||||
Db     398  CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy      723 CACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
          |||||
Db     458  CACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517

Qy      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842
          |||||
Db     518  CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAATGTCC 577

Qy      843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873
          ||| |  || |  |  || |  |  ||
Db     578  AAATGGATTCTTCGGACAGAGATGTTTGGAG 608
```

RESULT 10

AR116618

LOCUS AR116618 1476 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 7 from patent US 6133423.

ACCESSION AR116618

VERSION AR116618.1 GI:14096940

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1476)

AUTHORS Gearing,D.P. and Busfield,S.J.

TITLE Don-1 gene and polypeptides and uses therefor

JOURNAL Patent: US 6133423-A 7 17-OCT-2000;
 FEATURES Location/Qualifiers
 source 1. .1476
 /organism="unknown"
 BASE COUNT 335 a 473 c 452 g 216 t
 ORIGIN

Query Match 47.7%; Score 427.8; DB 6; Length 1476;
 Best Local Similarity 89.8%; Pred. No. 1.4e-71;
 Matches 459; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
          || || | || | | | | || | | | | | | | |
Db      98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
          |||||
Db      158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
          |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602
          |||||
Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 337

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
          |||||
Db      338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
          |||||
Db      398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy      723 CACCCTGTCATCCTGGTCTGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
          |||||
Db      458 CACCCTGTCATCCTGGTCTGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517

Qy      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842
          |||||
Db      518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAAATGTCC 577

Qy      843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873
          |||| | || | || || | ||
Db      578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608
  
```

RESULT 11

AR098155

LOCUS AR098155 2268 bp DNA linear PAT 14-FEB-2001
 DEFINITION Sequence 31 from patent US 6074841.
 ACCESSION AR098155
 VERSION AR098155.1 GI:12807412
 KEYWORDS .
 SOURCE Unknown.
 ORGANISM Unknown.

Unclassified.
 REFERENCE 1 (bases 1 to 2268)
 AUTHORS Gearing, D.P. and Busfield, S.J.
 TITLE Don-1 gene and polypeptides and uses therefor
 JOURNAL Patent: US 6074841-A 31 13-JUN-2000;
 FEATURES Location/Qualifiers
 source 1. .2268
 /organism="unknown"
 BASE COUNT 502 a 734 c 701 g 331 t
 ORIGIN

Query Match 47.7%; Score 427.8; DB 6; Length 2268;
 Best Local Similarity 89.8%; Pred. No. 1.4e-71;
 Matches 459; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
          || || | || | | | | | | | | | | | | | | | |
Db      98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
          |||||
Db      158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
          |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602
          |||||
Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 337

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
          |||||
Db      338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
          |||||
Db      398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy      723 CACCCTGTTCATCCTGGTTCGGGGCACGCGCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
          |||||
Db      458 CACCCTGTTCATCCTGGTTCGGGGCACGCGCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517

Qy      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842
          |||||
Db      518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577

Qy      843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873
          |||| | || | || | || |
Db      578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608
  
```

RESULT 12
 AR116627
 LOCUS AR116627 2268 bp DNA linear PAT 16-MAY-2001
 DEFINITION Sequence 31 from patent US 6133423.
 ACCESSION AR116627

VERSION AR116627.1 GI:14096949
 KEYWORDS .
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 2268)
 AUTHORS Gearing, D.P. and Busfield, S.J.
 TITLE Don-1 gene and polypeptides and uses therefor
 JOURNAL Patent: US 6133423-A 31 17-OCT-2000;
 FEATURES Location/Qualifiers
 source 1. .2268
 /organism="unknown"
 BASE COUNT 502 a 734 c 701 g 331 t
 ORIGIN

Query Match 47.7%; Score 427.8; DB 6; Length 2268;
 Best Local Similarity 89.8%; Pred. No. 1.4e-71;
 Matches 459; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy	363	CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG	422
Db	98	CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA	157
Qy	423	CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	482
Db	158	AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	217
Qy	483	ATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	542
Db	218	ATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	277
Qy	543	TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA	602
Db	278	TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA	337
Qy	603	CTCAGCACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	662
Db	338	CTCAGCACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	397
Qy	663	CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	722
Db	398	CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	457
Qy	723	CACCCTGTTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG	782
Db	458	CACCCTGTTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG	517
Qy	783	CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAAGTGTCC	842
Db	518	CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC	577
Qy	843	TGTGGGATACACCGGGGACAGGTGTCAGCAG	873
Db	578	AAATGGATTCTTCGGACAGAGATGTTTGGAG	608

RESULT 13

AB001576
 LOCUS AB001576 2188 bp mRNA linear ROD 13-FEB-1999
 DEFINITION Rattus sp. mRNA for NTAK alpha2-1p, partial cds.
 ACCESSION AB001576
 VERSION AB001576.1 GI:2605478
 KEYWORDS neural- and thymus-derived activator for ErbB kinases.
 SOURCE Rattus sp.
 ORGANISM Rattus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (sites)
 AUTHORS Higashiyama,S., Horikawa,M., Yamada,K., Ichino,N., Nakano,N.,
 Nakagawa,T., Miyagawa,J., Matsushita,N., Nagatsu,T., Taniguchi,N.
 and Ishiguro,H.
 TITLE A novel brain-derived member of the epidermal growth factor family
 that interacts with ErbB3 and ErbB4
 JOURNAL J. Biochem. 122 (3), 675-680 (1997)
 MEDLINE 98006324
 PUBMED 9348101
 REFERENCE 2 (bases 1 to 2188)
 AUTHORS Ishiguro,H.
 TITLE Direct Submission
 JOURNAL Submitted (04-MAR-1997) Hiroshi Ishiguro, Fujita Health University,
 ICMS; 1-98, kutsukake-cho, Toyoake, Aichi 470-11, Japan
 (E-mail:hishi@fujita-hu.ac.jp, Tel:0562-93-9393, Fax:0562-93-8831)
 FEATURES
 source Location/Qualifiers
 1. .2188
 /organism="Rattus sp."
 /mol_type="mRNA"
 /db_xref="taxon:10118"
 /cell_line="PC12"
 /cell_type="pheochromocytoma"
 CDS
 <1. .1926
 /note="neural- and thymus-derived activator for ErbB
 kinases (NTAK); a member of the epidermal growth factor
 (EGF) family"
 /codon_start=1
 /product="NTAK alpha2-1p"
 /protein_id="BAA23348.1"
 /db_xref="GI:2605479"
 /translation="FFFFKTAFAVPDPNGKNIKKEVGKILCTDCATRPKLKKMKSQTG
 EVGEKQSLKCEAAAGNPQPSYRWFKDGKELNRSRDIRIKYGNRKN SRLQFNKV KVED
 AGEYVCEAENILGKDTVGRHLVNSVSTTLSSWSGHARKCNETAKSYCVNGGVCYYIE
 GINQLSCKCPNGFFGQRCLEKLPLRLYMPDPKQKAEELYQKRVLTITGICVALLVVGI
 VCVVAYCKTKKQRRQMHHHLRQNMCPAHQNRSLANGPSHPRLDPEEIQMADYISKVNP
 ATDHVIRREAETTFSGSHSCSPSHHCSTATPTSSHRHESHTWSLERSESITSDSQSGI
 MLSSVGTSKCN SPACVEARARRAAAYSQEERRRAAMPPYHDSIDSLRDSPHSERYVSA
 LTPARLSPVDFHYSLATQVPTFEITSPNSAHAVSLPPAAPISYRLAEQQPLLRHPAP
 PGPGPGPGADMQRSYDSYYPAGPGPRRGACALGGSGLGSLPASPFRIPEDEYETQ
 ECAPPPPPRPRTRGASRRTSAGPRRWRRSRLNGLAAQRARAARDSLSLSSGSGCGSAS
 ASDDDADDADGALAAESTPFLGLRAAHDALRSDSPPLCPAADSRTYYSLDSHSTRASS
 RHSRGPPTRAKQDSGPL"
 BASE COUNT 515 a 674 c 643 g 356 t
 ORIGIN
 Query Match 47.4%; Score 425; DB 10; Length 2188;

Best Local Similarity 87.0%; Pred. No. 4.7e-71;
Matches 467; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```

QY      337 TTAGTCTTTAAGACGGCCTTTGCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAG 396
      || | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      4 TTTTTTTTAAAGACAGCCTTTGCCCCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAG 63

QY      397 GTGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGC 456
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      64 GTGGGCAAGATCCTGTGCACTGACTGCGCAACCCGGCCCAAGCTGAAGAAGATGAAGAGT 123

QY      457 CAGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCC 516
      ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      124 CAGACAGGAGAGGTGGGCGAGAAGCAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCC 183

QY      517 CAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGC 576
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db      184 CAGCCCTCCTATCGATGGTTCAAGGACGGCAAGGAGCTCAACCGGAGTCGTGACATTTCGC 243

QY      577 ATCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAG 636
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      244 ATCAAGTATGGCAACGGCAGAAAGAACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAG 303

QY      637 GACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGC 696
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      304 GACGCTGGAGAGTACGTCTGTGAGGCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGC 363

QY      697 CGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCACTCCTGGTTCGGGGCACGCCCCGGAAG 756
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      364 CGGCTCCATGTCAACAGTGTGAGCACCCTGTGTCGTCCTGGTTCGGGGCACGCCCCGGAAG 423

QY      757 TGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGC 816
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      424 TGCAATGAGACAGCCAAGTCCTACTGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGC 483

QY      817 ATCAACCAGCTCTCCTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      484 ATCAACCAACTCTCCTGCAAATGTCCAAACGGATTCTTCGGACAGAGATGTTTGGAG 540

```

RESULT 14

AC094080

LOCUS

AC094080

118504 bp

DNA

linear

PRI 27-MAR-2002

DEFINITION

Homo sapiens chromosome 5 clone CTB-77K22, complete sequence.

ACCESSION

AC094080

VERSION

AC094080.4 GI:19747152

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 118504)

AUTHORS

DOE Joint Genome Institute and Stanford Human Genome Center.

TITLE

Direct Submission

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 118504)

AUTHORS

DOE Joint Genome Institute.

TITLE Direct Submission
 JOURNAL Submitted (14-SEP-2001) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 3 (bases 1 to 118504)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 4 (bases 1 to 118504)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (27-MAR-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 COMMENT On Mar 27, 2002 this sequence version replaced gi:19224838.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 Finishing Completed at Stanford Human Genome Center
 www-shgc.stanford.edu
 Quality: Phrap Quality >=40 99.7% of Sequence;
 Estimated Total Number of Errors is 0.5.
 FEATURES Location/Qualifiers
 source 1. .118504
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="CTB-77K22"
 BASE COUNT 33986 a 23829 c 24682 g 36007 t
 ORIGIN

Query Match 47.4%; Score 424.8; DB 9; Length 118504;
 Best Local Similarity 94.2%; Pred. No. 3.9e-71;
 Matches 452; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

QY 1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 80890 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 80949
 QY 61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 80950 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 81009
 QY 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 81010 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 81069
 QY 181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 81070 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 81129
 QY 241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 81130 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 81189
 QY 301 CGCTACATCTTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 81190 CGCTACATCTTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 81249

Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 147077 bases at least Q40
 Consensus quality: 149570 bases at least Q30
 Consensus quality: 150764 bases at least Q20
 Insert size: 188000; agarose-fp
 Insert size: 151938; sum-of-contigs
 Quality coverage: 4.4 in Q20 bases; agarose-fp
 Quality coverage: 5.4 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 10 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1840: contig of 1840 bp in length
 * 1841 1940: gap of 100 bp
 * 1941 5083: contig of 3143 bp in length
 * 5084 5183: gap of 100 bp
 * 5184 6473: contig of 1290 bp in length
 * 6474 6573: gap of 100 bp
 * 6574 11402: contig of 4829 bp in length
 * 11403 11502: gap of 100 bp
 * 11503 25453: contig of 13951 bp in length
 * 25454 25553: gap of 100 bp
 * 25554 40778: contig of 15225 bp in length
 * 40779 40878: gap of 100 bp
 * 40879 58024: contig of 17146 bp in length
 * 58025 58124: gap of 100 bp
 * 58125 87982: contig of 29858 bp in length
 * 87983 88082: gap of 100 bp
 * 88083 121177: contig of 33095 bp in length
 * 121178 121277: gap of 100 bp
 * 121278 152838: contig of 31561 bp in length.

FEATURES	Location/Qualifiers
source	1. .152838
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/clone="RP11-13018"
	/clone_lib="RPC1-11 Human Male BAC"
misc_feature	1. .1840
	/note="assembly_fragment"
misc_feature	1941. .5083
	/note="assembly_fragment"
misc_feature	5184. .6473
	/note="assembly_fragment"
	clone_end:SP6
	vector_side:right"
misc_feature	6574. .11402
	/note="assembly_fragment"
misc_feature	11503. .25453
	/note="assembly_fragment"

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2004, 07:12:21 ; Search time 290.304 Seconds
(without alignments)
8340.911 Million cell updates/sec

Title: US-09-864-675-3
Perfect score: 897
Sequence: 1 atgaggcgcgacccggcccc.....caatgggtcaacttctcctaa 897

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	% Query		Match	Length	DB	ID	Description
	Score	Match					
1	897	100.0	897	24	AAS18020	Human cDNA encodin	
2	849	94.6	994	24	AAS18019	Human cDNA encodin	
3	835.4	93.1	1884	19	AAV17814	Homo sapiens don-1	
4	784	87.4	3441	18	AAT87922	Rat cerebellum der	
5	738.6	82.3	1803	18	AAT87923	Rat cerebellum der	
6	737	82.2	3076	19	AAV43674	Receptor type tyro	
7	723.4	80.6	1863	25	ABS56035	cDNA encoding huma	
8	427.8	47.7	2268	19	AAV17816	Homo sapiens don-1	
9	426.8	47.6	1474	25	ABS56036	cDNA encoding huma	
10	426.8	47.6	2266	25	ABS56045	cDNA encoding huma	
11	426.2	47.5	1476	19	AAV17815	Homo sapiens don-1	
12	424	47.3	1054	24	ABL40993	Human neuregulin 2	
13	405.4	45.2	1607	19	AAV17813	Mus musculus don-1	
14	402.2	44.8	2467	19	AAV17812	Mus musculus don-1	
15	393.4	43.9	1561	25	ABS56034	cDNA encoding muri	
16	393.4	43.9	2442	25	ABS56033	cDNA encoding muri	
17	256	28.5	667	18	AAT87924	Human cerebellum d	
18	173	19.3	419	24	ABL40994	Human neuregulin 2	
19	124.6	13.9	480	24	ABL40995	Human neuregulin 2	
20	124	13.8	350	24	ABL40996	Human neuregulin 2	
21	85.6	9.5	1986	20	AAZ32061	Human METH2 relate	
22	84	9.4	1986	22	AAC90318	L12260 cDNA clone.	
23	84	9.4	2003	17	AAT48090	Human neuregulin G	
24	84	9.4	2003	17	AAT30995	Glial growth facto	
25	84	9.4	2003	17	AAT06739	Glial growth facto	
26	84	9.4	2003	20	AAZ32062	Human METH2 relate	
27	84	9.4	2003	22	AAC90319	I36352 cDNA clone.	
28	83.4	9.3	1108	13	AAQ30671	GGF2BPP3.CDS. Syn	
29	83.4	9.3	1108	15	AAQ58304	GGF-II cDNA sequen	
30	83.4	9.3	1108	15	AAQ62850	GGF-II cDNA sequen	
31	83.4	9.3	1108	16	AAQ74886	Putative bovine gl	
32	83.4	9.3	1108	17	AAT48080	Bovine neuregulin	
33	83.4	9.3	1108	17	AAT30998	Bovine glial growt	
34	83.4	9.3	1108	17	AAT06704	Bovine glial growt	
35	82.4	9.2	2003	15	AAQ62838	GGF2HBS5. Homo sa	
36	82.4	9.2	2003	15	AAQ58329	GGF2HBS5. Homo sa	
37	78.2	8.7	1027	22	AAF80062	Nucleotide sequenc	
38	78.2	8.7	3086	22	AAF80059	Nucleotide sequenc	
39	76.8	8.6	1140	15	AAQ58321	GGF2BPP2. Bos tau	
40	76.8	8.6	1140	15	AAQ62840	GGF2BPP2. Bos tau	
41	76.8	8.6	1140	16	AAQ74912	Bovine glial cell	
42	76.8	8.6	1140	17	AAT48088	Human neuregulin G	
43	76.8	8.6	1140	17	AAT31001	Glial growth facto	
44	76.8	8.6	1140	17	AAT06731	BPP2 glial growth	
45	76.8	8.6	1140	20	AAX81201	Nucleotide sequenc	

ALIGNMENTS

RESULT 1

AAS18020

ID AAS18020 standard; cDNA; 897 BP.

XX

AC AAS18020;

XX

DT 12-MAR-2002 (first entry)

XX

DE Human cDNA encoding Neuregulin-2beta, NRG-2beta.

XX

KW Human; ss; neuregulin-2; NRG-2alpha; NRG-2beta; mitogenesis;
 KW cell survival; cell growth; cell differentiation; erbB receptor;
 KW cardiomyopathy; ischaemic damage; cardiac trauma; heart failure;
 KW atherosclerosis; vascular lesion; vascular hypertension;
 KW degenerative congenital vascular disease; myasthenia gravis;
 KW neurodegenerative disorder; peripheral neuropathy;
 KW sensory nerve fiber neuropathy; motor fiber neuropathy;
 KW sensory nerve fiber neuropathy; multiple sclerosis;
 KW amyotrophic lateral sclerosis; spinal muscular atrophy; nerve injury;
 KW Alzheimer's disease; Parkinson's disease; cerebellar ataxia;
 KW spinal cord injury; tumour; neurofibromatosis; transgenic animal.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
FT	CDS	1..897
FT		/*tag= a
FT		/product= "NRG-2beta"

XX

PN WO200189568-A1.

XX

PD 29-NOV-2001.

XX

PF 23-MAY-2001; 2001WO-US16896.

XX

PR 23-MAY-2000; 2000US-206495P.

XX

PA (CENE-) CENES PHARM INC.

XX

PI Marchionni MA;

XX

DR WPI; 2002-097612/13.

DR P-PSDB; AAU11636.

XX

PT Neuregulin-2 polypeptide and polynucleotide useful for treating
 PT multiple sclerosis, spinal muscular atrophy, nerve injury, Alzheimer's
 PT disease, by increasing mitogenesis, survival, growth or differentiation
 PT of a cell -

XX

PS Claim 57; Fig 8; 79pp; English.

XX

CC The invention relates to a substantially pure neuregulin (NRG)-2
 CC polypeptide comprising or consisting of a sequence for human
 CC NRG-2alpha or NRG-2beta (clone 2b7) and the polynucleotides encoding
 CC the. Also included are a vector expressing the protein, a host cell
 CC comprising the vector, a transgenic non-human animal transformed with

CC the vector or having a knockout mutation in one or both NRG-2
 CC alleles and an anti-NRG-2 antibody. Analysis of mutations in NRG-2 in an
 CC individual is useful for diagnosing an increased likelihood of
 CC developing a NRG-2-related disease or condition in a test subject.
 CC NRG-2 is useful for increasing the mitogenesis, survival, growth or
 CC differentiation of a cell (e.g. a neuronal cell), where the cell
 CC expresses an erbB receptor. NRG-2 is useful for treating diseases
 CC and disorders such as cardiomyopathy (preferably degenerative congenital
 CC disease), ischaemic damage, cardiac trauma or heart failure or which
 CC has a condition affecting smooth muscle which include atherosclerosis,
 CC vascular lesion, vascular hypertension, and degenerative congenital
 CC vascular disease, myasthenia gravis, a neurodegenerative disorder,
 CC peripheral neuropathy, a sensory nerve fiber neuropathy, a motor fiber
 CC and a sensory nerve fiber neuropathy, multiple sclerosis, amyotrophic
 CC lateral sclerosis, spinal muscular atrophy, nerve injury, Alzheimer's
 CC disease, Parkinson's disease, cerebellar ataxia, and spinal cord injury.
 CC The antibody is useful for treatment of a tumour comprising inhibiting
 CC proliferation of a tumour cell preferably a glial tumour cell, for
 CC treating of neurofibromatosis by inhibiting glial cell mitogenesis.
 CC The present sequence encodes NRG-2beta.

XX

SQ Sequence 897 BP; 200 A; 261 C; 282 G; 154 T; 0 other;

Query Match 100.0%; Score 897; DB 24; Length 897;
 Best Local Similarity 100.0%; Pred. No. 2.4e-199;
 Matches 897; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC	60
Db	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC	60
Qy	61	TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	61	TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Qy	301	CGCTACATCTTTTCTCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	301	CGCTACATCTTTTCTCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480

Db 421 TGC GCCACCCG G C C C AAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480

Qy 481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
 |||

Db 481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540

Qy 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 600
 |||

Db 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 600

Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
 |||

Db 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660

Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
 |||

Db 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720

Qy 721 ACCACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
 |||

Db 721 ACCACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780

Qy 781 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
 |||

Db 781 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840

Qy 841 CCTGTGGGATACACCGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTCCTAA 897
 |||

Db 841 CCTGTGGGATACACCGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTCCTAA 897

RESULT 2

AAS18019

ID AAS18019 standard; cDNA; 994 BP.

XX

AC AAS18019;

XX

DT 12-MAR-2002 (first entry)

XX

DE Human cDNA encoding Neuregulin-2alpha, NRG-2alpha.

XX

KW Human; ss; neuregulin-2; NRG-2alpha; NRG-2beta; mitogenesis;
 KW cell survival; cell growth; cell differentiation; erbB receptor;
 KW cardiomyopathy; ischaemic damage; cardiac trauma; heart failure;
 KW atherosclerosis; vascular lesion; vascular hypertension;
 KW degenerative congenital vascular disease; myasthenia gravis;
 KW neurodegenerative disorder; peripheral neuropathy;
 KW sensory nerve fiber neuropathy; motor fiber neuropathy;
 KW sensory nerve fiber neuropathy; multiple sclerosis;
 KW amyotrophic lateral sclerosis; spinal muscular atrophy; nerve injury;
 KW Alzheimer's disease; Parkinson's disease; cerebellar ataxia;
 KW spinal cord injury; tumour; neurofibromatosis; transgenic animal.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..993

```

FT          /*tag= a
FT          /product= "NRG-2alpha"
XX
PN  WO200189568-A1.
XX
PD  29-NOV-2001.
XX
PF  23-MAY-2001; 2001WO-US16896.
XX
PR  23-MAY-2000; 2000US-206495P.
XX
PA  (CENE-) CENES PHARM INC.
XX
PI  Marchionni MA;
XX
DR  WPI; 2002-097612/13.
DR  P-PSDB; AAU11635.
XX
PT  Neuregulin-2 polypeptide and polynucleotide useful for treating
PT  multiple sclerosis, spinal muscular atrophy, nerve injury, Alzheimer's
PT  disease, by increasing mitogenesis, survival, growth or differentiation
PT  of a cell -
XX
PS  Claim 57; Fig 6; 79pp; English.
XX
CC  The invention relates to a substantially pure neuregulin (NRG)-2
CC  polypeptide comprising or consisting of a sequence for human
CC  NRG-2alpha or NRG-2beta (clone 2b7) and the polynucleotides encoding
CC  the. Also included are a vector expressing the protein, a host cell
CC  comprising the vector, a transgenic non-human animal transformed with
CC  the vector or having a knockout mutation in one or both NRG-2
CC  alleles and an anti-NRG-2 antibody. Analysis of mutations in NRG-2 in an
CC  individual is useful for diagnosing an increased likelihood of
CC  developing a NRG-2-related disease or condition in a test subject.
CC  NRG-2 is useful for increasing the mitogenesis, survival, growth or
CC  differentiation of a cell (e.g. a neuronal cell), where the cell
CC  expresses an erbB receptor. NRG-2 is useful for treating diseases
CC  and disorders such as cardiomyopathy (preferably degenerative congenital
CC  disease), ischaemic damage, cardiac trauma or heart failure or which
CC  has a condition affecting smooth muscle which include atherosclerosis,
CC  vascular lesion, vascular hypertension, and degenerative congenital
CC  vascular disease, myasthenia gravis, a neurodegenerative disorder,
CC  peripheral neuropathy, a sensory nerve fiber neuropathy, a motor fiber
CC  and a sensory nerve fiber neuropathy, multiple sclerosis, amyotrophic
CC  lateral sclerosis, spinal muscular atrophy, nerve injury, Alzheimer's
CC  disease, Parkinson's disease, cerebellar ataxia, and spinal cord injury.
CC  The antibody is useful for treatment of a tumour comprising inhibiting
CC  proliferation of a tumour cell preferably a glial tumour cell, for
CC  treating of neurofibromatosis by inhibiting glial cell mitogenesis.
CC  The present sequence encodes NRG-2alpha.
XX
SQ  Sequence 994 BP; 230 A; 279 C; 304 G; 181 T; 0 other;

Query Match          94.6%; Score 849; DB 24; Length 994;
Best Local Similarity 98.3%; Pred. No. 3.6e-188;
Matches 858; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC	60
Db	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC	60
Qy	61	TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	61	TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Qy	301	CGCTACATCTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	301	CGCTACATCTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG	600
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Qy	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT	840
Db	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT	840
Qy	841	CCTGTGGGATACACCGGGGACAGGTGTCAGCAG	873

Db || |||| | || | || ||| | ||
841 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 873

RESULT 3

AAV17814

ID AAV17814 standard; cDNA; 1884 BP.

XX

AC AAV17814;

XX

DT 17-AUG-1998 (first entry)

XX

DE Homo sapiens don-1 gene splice variant.

XX

KW Murine; don-1 gene; melanoma; treatment; adenocarcinoma;
KW epithelial cell; proliferation; stimulation; treatment; tumours;
KW skin; oesophagus; lung; breast; liver; pancreas; colon; prostate;
KW gastrointestinal tract; uterus; wound healing; transmembrane; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 664..1884

FT /*tag= a

FT /note= "don-1 polypeptide"

XX

PN WO9807736-A1.

XX

PD 26-FEB-1998.

XX

PF 18-AUG-1997; 97WO-US14585.

XX

PR 19-NOV-1996; 96US-0753007.

PR 19-AUG-1996; 96US-0699591.

XX

PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.

XX

PI Busfield SJ, Gearing DP;

XX

DR WPI; 1998-169084/15.

DR P-PSDB; AAW48381.

XX

PT Mouse and human don-1 polypeptide(s) - useful for treatment of
PT melanomas and adenocarcinoma(s), and for wound healing

XX

PS Claim 4; Fig 3; 121pp; English.

XX

CC The sequence is that of a human don-1 gene splice variant.
CC Don-1 polypeptides stimulate proliferation of epithelial cells
CC and thus are implicated in melanomas and adenocarcinomas in which
CC epithelial cells proliferate out of control. Compounds that
CC interfere with don-1 mediated cell proliferation can be used
CC in the treatment of tumours such as melanomas and adenocarcinomas
CC of the skin, oesophagus, lung, breast, liver, pancreas,
CC gastrointestinal tract, colon, prostate or uterus. Alternatively,
CC don-1 polypeptides can be used to stimulate epithelial cell
CC proliferation, e.g. for wound healing.

XX

SQ Sequence 1884 BP; 426 A; 607 C; 560 G; 291 T; 0 other;

Query Match 93.1%; Score 835.4; DB 19; Length 1884;
Best Local Similarity 98.1%; Pred. No. 6.2e-185;
Matches 856; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

```
QY      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTGC 60
      |||
Db      218 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTGC 277

QY      61 TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
      |||
Db      278 TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 337

QY     121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
      |||
Db     338 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 397

QY     181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
      |||
Db     398 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 457

QY     241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
      |||
Db     458 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 517

QY     301 CGCTACATCTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
      |||
Db     518 CGCTACATCTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 577

QY     361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
      |||
Db     578 CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC 636

QY     421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
      |||
Db     637 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 696

QY     481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
      |||
Db     697 CAATCGCTGAAGTGTGAGGCAGCAGCCGGAATCCCCAGCCTTCCTACCGTTGGTTCAAG 756

QY     541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG 600
      |||
Db     757 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG 816

QY     601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
      |||
Db     817 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 876

QY     661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
      |||
Db     877 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 936

QY     721 ACCACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
      |||
```

Db 937 ACCACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 996
 QY 781 TGGCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
 |||||
 Db 997 TGGCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 1056
 QY 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873
 || |||| | || | || || ||
 Db 1057 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 1089

RESULT 4

AAT87922

ID AAT87922 standard; cDNA; 3441 BP.

XX

AC AAT87922;

XX

DT 18-DEC-1997 (first entry)

XX

DE Rat cerebellum derived growth factor 1 cDNA.

XX

KW Rat; cerebellum derived growth factor; CDGF1; screening; binding;
 KW modulation; erbB type receptor; identification; indication; risk;
 KW proliferation; differentiation; induction; neuron; hyperplasia;
 KW stem cell culture; intracerebral graft; alleviation; repair;
 KW behavioural defect; nervous system; central; peripheral; nerve;
 KW prosthesis; damage; entubulation; cell survival; treatment;
 KW injury; trauma; ischaemia; ischemia; stroke; infection; disorder;
 KW inflammation; neurodegeneration; disease; Parkinson's;
 KW Huntingdon's; amyotrophic lateral sclerosis; sensory; retina;
 KW spinocerebellar degeneration; multiple sclerosis; neoplasia;
 KW amalignant glioma; medulloblastoma; neuroectodermal tumour; ds.

XX

OS Rattus rattus.

XX

FH Key Location/Qualifiers

FT CDS 180..2444

FT /*tag= a

FT sig_peptide 180..248

FT /*tag= b

FT mat_peptide 249..2441

FT /*tag= c

FT /product= cerebellum_derived_growth_factor

XX

PN W09709425-A1.

XX

PD 13-MAR-1997.

XX

PF 09-SEP-1996; 96WO-US14484.

XX

PR 08-SEP-1995; 95US-0525864.

XX

PA (HARD) HARVARD COLLEGE.

PA (STRD) UNIV LELAND STANFORD JUNIOR.

PA (STRD) UNIV LELAND S STANFORD.

XX

PI Chang H;

XX
 DR WPI; 1997-192900/17.
 DR P-PSDB; AAW27536.
 XX
 PT Rat and human cerebellum-derived growth factors - used in the
 PT treatment of neuronal injury and proliferative disorders
 XX
 PS Claim 17; Pages 63-66; 94pp; English.
 XX
 CC The present sequence encodes rat cerebellum derived growth factor 1
 CC (CDGF1), which can be used to screen for modulators of CDGF
 CC binding to erbB type receptors. Identification of a modification or
 CC mutation in a CDGF gene, or aberrant expression of a CDGF gene or
 CC levels of soluble CDGF may be used to indicate the risk of unwanted
 CC cell proliferation or differentiation.
 CC CDGF may be used to induce neuronal differentiation in stem cell
 CC culture, and maintain the integrity of a terminally differentiated
 CC neuronal cell culture, e.g. useful for intracerebral grafting to
 CC alleviate behavioural defects. CDGF may also be used in nerve
 CC protheses to repair central and peripheral nerve damage, especially
 CC where a crushed or severed axon is entubulated by a prosthetic.
 CC CDGF may also be used to enhance neuronal cell survival in the
 CC central or peripheral nervous system, to treat neurological
 CC conditions associated with nervous system injury, e.g. traumatic,
 CC chemical or vasal injury and deficits such as ischaemia resulting
 CC from stroke, infectious/inflammatory and tumour induced injury,
 CC chronic neurodegenerative disease including Parkinson's and
 CC Huntingdon's, amyotrophic lateral sclerosis, spinocerebellar
 CC degeneration, chronic immunological disease of the nervous system
 CC including multiple sclerosis, disorders of the sensory neurons and
 CC degenerative diseases of the retina. CDGF may also be used to treat
 CC neoplastic or hyperplastic transformations, particularly of the
 CC central nervous system, e.g. amalignant gliomas, medulloblastomas
 CC and neuroectodermal tumours.
 XX
 SQ Sequence 3441 BP; 777 A; 1057 C; 1015 G; 592 T; 0 other;

Query Match 87.4%; Score 784; DB 18; Length 3441;
 Best Local Similarity 92.2%; Pred. No. 6.7e-173;
 Matches 826; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCTCGCTCGCCTGC	60
Db	180	ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC	239
Qy	61	TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	240	TACTCGCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	299
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	300	GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTCCAGCTCTAACAGCACCCGAGAGCCT	359
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	360	CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	419

Qy 241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
 |||
 Db 420 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 479
 Qy 301 CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
 |||
 Db 480 CGCTACATCTTTTTCCTGGAGCCCACCGAGCAGCCCTTAGTCTTTAAGACAGCCTTTGCC 539
 Qy 361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
 || ||| | |||
 Db 540 CCGGTCGACCCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 599
 Qy 421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
 ||| | |||
 Db 600 TGCGCAACCCGGCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG 659
 Qy 481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
 || ||| | ||| || || || |||
 Db 660 CAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCCCAGCCCTCCTATCGATGGTTCAAG 719
 Qy 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 600
 || |||
 Db 720 GACGGCAAGGAGCTCAACCGGAGTCGTGACATTTCGCATCAAGTATGGCAACGGCAGAAAG 779
 Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
 |||
 Db 780 AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG 839
 Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
 || |||
 Db 840 GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC 899
 Qy 721 ACCACCCTGTCTCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
 |||
 Db 900 ACCACTCTGTCTCCTGGTTCGGGGCACGCCCCGGAAGTGCAATGAGACAGCCAAGTCCTAC 959
 Qy 781 TGCCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
 || |||
 Db 960 TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAAGTGT 1019
 Qy 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTTCGCAATGGTCAACTTCTCCTA 896
 |||
 Db 1020 CCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTTCGCAATGGTCAACTTCTCCAA 1075

RESULT 5

AAT87923

ID AAT87923 standard; cDNA; 1803 BP.

XX

AC AAT87923;

XX

DT 18-DEC-1997 (first entry)

XX

DE Rat cerebellum derived growth factor 2 cDNA.

XX

KW Rat; cerebellum derived growth factor; CDGF2; screening; binding;

KW modulation; erbB type receptor; identification; indication; risk;

KW proliferation; differentiation; induction; neuron; hyperplasia;
 KW stem cell culture; intracerebral graft; alleviation; repair;
 KW behavioural defect; nervous system; central; peripheral; nerve;
 KW prosthesis; damage; entubulation; cell survival; treatment;
 KW injury; trauma; ischaemia; ischemia; stroke; infection; disorder;
 KW inflammation; neurodegeneration; disease; Parkinson's;
 KW Huntingdon's; amyotrophic lateral sclerosis; sensory; retina;
 KW spinocerebellar degeneration; multiple sclerosis; neoplasia;
 KW amalignant glioma; medulloblastoma; neuroectodermal tumour; ds.
 XX
 OS Rattus rattus.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..993
 FT /*tag= a
 FT sig_peptide 1..69
 FT /*tag= b
 FT mat_peptide 70..990
 FT /*tag= c
 FT /product= cerebellum_derived_growth_factor
 XX
 PN WO9709425-A1.
 XX
 PD 13-MAR-1997.
 XX
 PF 09-SEP-1996; 96WO-US14484.
 XX
 PR 08-SEP-1995; 95US-0525864.
 XX
 PA (HARD) HARVARD COLLEGE.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PA (STRD) UNIV LELAND S STANFORD.
 XX
 PI Chang H;
 XX
 DR WPI; 1997-192900/17.
 DR P-PSDB; AAW27537.
 XX
 PT Rat and human cerebellum-derived growth factors - used in the
 PT treatment of neuronal injury and proliferative disorders
 XX
 PS Claim 17; Pages 70-74; 94pp; English.
 XX
 CC The present sequence encodes rat cerebellum derived growth factor 2
 CC (CDGF2), which can be used to screen for modulators of CDGF
 CC binding to erbB type receptors. Identification of a modification or
 CC mutation in a CDGF gene, or aberrant expression of a CDGF gene or
 CC levels of soluble CDGF may be used to indicate the risk of unwanted
 CC cell proliferation or differentiation.
 CC CDGF may be used to induce neuronal differentiation in stem cell
 CC culture, and maintain the integrity of a terminally differentiated
 CC neuronal cell culture, e.g. useful for intracerebral grafting to
 CC alleviate behavioural defects. CDGF may also be used in nerve
 CC prostheses to repair central and peripheral nerve damage, especially
 CC where a crushed or severed axon is entubulated by a prosthetic.
 CC CDGF may also be used to enhance neuronal cell survival in the
 CC central or peripheral nervous system, to treat neurological

CC conditions associated with nervous system injury, e.g. traumatic,
 CC chemical or vasal injury and deficits such as ischaemia resulting
 CC from stroke, infectious/inflammatory and tumour induced injury,
 CC chronic neurodegenerative disease including Parkinson's and
 CC Huntingdon's, amyotrophic lateral sclerosis, spinocerebellar
 CC degeneration, chronic immunological disease of the nervous system
 CC including multiple sclerosis, disorders of the sensory neurons and
 CC degenerative diseases of the retina. CDGF may also be used to treat
 CC neoplastic or hyperplastic transformations, particularly of the
 CC central nervous system, e.g. amalignant gliomas, medulloblastomas
 CC and neuroectodermal tumours.

XX

SQ Sequence 1803 BP; 408 A; 549 C; 537 G; 309 T; 0 other;

Query Match 82.3%; Score 738.6; DB 18; Length 1803;
 Best Local Similarity 90.4%; Pred. No. 2.2e-162;
 Matches 789; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTGC	60
Db	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC	60
Qy	61	TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	61	TACTCGCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	121	GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT	180
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	181	CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG	300
Qy	301	CGCTACATCTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	301	CGCTACATCTTTTCTGGAGCCACCGAGCAGCCCTTAGTCTTTAAGACAGCCTTTGCC	360
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	361	CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	421	TGCGCAACCCGGCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG	480
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	481	CAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCCCAGCCTCCTATCGATGGTTCAAG	540
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAAG	600
Db	541	GACGGCAAGGAGCTCAACCGGAGTCGTGACATTTCGCATCAAGTATGGCAACGGCAGAAAAG	600

XX

PS Examples; Pages 9-13; 17pp; Japanese.

XX

CC This cDNA encodes the ligand of receptor type tyrosine kinase ErbB. A
CC prokaryotic or eukaryotic host cell transformed by a recombinant vector
CC containing the encoding DNA can be used for the recombinant production of
CC the protein. The invention provides a method for inhibiting the formation
CC of the ligand of receptor type tyrosine kinase ErbB in an animal using
CC an antibody recognizing the protein. The ligand of the tyrosine kinase
CC ErbB receptor and associated materials can be used for treating or
CC diagnosing nervous diseases and cancers.

XX

SQ Sequence 3076 BP; 673 A; 996 C; 944 G; 463 T; 0 other;

Query Match 82.2%; Score 737; DB 19; Length 3076;

Best Local Similarity 90.3%; Pred. No. 5.8e-162;

Matches 788; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```
QY      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCTCGCTCGCCTGC 60
      |||
Db      556 ATGAGGCGCGACCCGGCCCCCGGCTCCTCGATGCTGCTCTTCGGTGTGTCTACTCGCCTGC 615

QY      61 TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
      |||
Db      616 TACTCGCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 675

QY     121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
      |||
Db      676 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT 735

QY     181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
      |||
Db      736 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 795

QY     241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
      |||
Db      796 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 855

QY     301 CGCTACATCTTTTCTGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
      |||
Db      856 CGCTACATCTTTTCTGAGCCACCGAGCAGCCCTTAGTCTTTAAGACAGCCTTTGCC 915

QY     361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
      ||
Db      916 CCGGTGACCCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 975

QY     421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
      |||
Db      976 TGCGCAACCCGGCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG 1035

QY     481 CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
      ||
Db     1036 CAGTCGCTCAAGTGTGAGGCGGCGGCGGGGAACCCCCAGCCCTCCTATCGATGGTTCAAG 1095

QY     541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 600
      ||
Db     1096 GACGGCAAGGAGCTCAACCGGAGTCGTGACATTTCGCATCAAGTATGGCAACGGCAGAAAG 1155
```


XX
 PI Gearing DP, Busfield SJ;
 XX
 DR WPI; 2003-039584/03.
 DR P-PSDB; ABG71638.
 XX
 PT Novel Don-1 polypeptide useful for stimulating proliferation of cells,
 PT for identifying proteins that interact with Don-1, and for regulating
 PT tumour formation and progression in brain -
 XX
 PS Claim 4; Fig 3; 66pp; English.
 XX

CC The present invention relates to the isolation of a novel gene
 CC called Don-1, and alternate splice variants of Don-1, which are
 CC related to epidermal growth factors (EGF) such as neuregulins.
 CC Don-1 polypeptides are glycoprotein ligands. Both murine and human
 CC Don-1 sequences are cloned. The mouse Don-1 gene maps to chromosome 18.
 CC Don-1 polypeptides are useful for stimulating proliferation of a cell.
 CC Antibodies to Don-1 polypeptides are useful for detecting Don-1
 CC in a sample. The Don-1 polypeptides are useful for treating and
 CC diagnosing cell proliferative disorders and play a role in the
 CC proliferation of carcinomas e.g. adenocarcinoma, myeloma, in cell
 CC differentiation, proliferation and survival. The polypeptides are
 CC also useful for inhibiting proliferation of adenocarcinoma cells,
 CC for stimulating the proliferation of cells such as epithelial cells
 CC to promote wound healing, for identifying proteins that interact
 CC with Don-1, and for regulating tumour formation and progression in
 CC the brain. The polynucleotide sequences encoding Don-1 may be used
 CC in gene therapy. The present sequence encodes human membrane-bound
 CC splice variant of Don-1.
 XX

SQ Sequence 1863 BP; 422 A; 602 C; 553 G; 286 T; 0 other;

Query Match 80.6%; Score 723.4; DB 25; Length 1863;
 Best Local Similarity 96.2%; Pred. No. 7.5e-159;
 Matches 840; Conservative 0; Mismatches 16; Indels 17; Gaps 9;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC	60
Db	213	ATGAGGCGCGACCCGGCCCC--CTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC	270
Qy	61	TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	271	TACTCGCCAGCCTCAAGTCA--GCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	328
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	329	GGCAAGGTACAGGGGCTGGT--CAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	386
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGCCGCTCCGGAGCGGG	240
Db	387	CCCGCCTCGGGTCGGGTGGCG--GGTAAAGGTGCTGGACAAGTGCCGCTCCGGAGCGGG	444
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	445	GGGCTGCAGCGCGAGCAGGTG--CAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	502

QY 301 CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
 |||
 Db 503 CGCTACATCTTTTTCCTGGAG--CACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 560
 QY 361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
 |||
 Db 561 CCCCT-GATACCAACGGCAAAA--CTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC 617
 QY 421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
 |||
 Db 618 TGCGCCACCCGGCCCAAGTTGA--AAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 675
 QY 481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
 |||
 Db 676 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 735
 QY 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 600
 |||
 Db 736 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 795
 QY 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
 |||
 Db 796 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 855
 QY 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
 |||
 Db 856 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 915
 QY 721 ACCACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
 |||
 Db 916 ACCACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 975
 QY 781 TGCCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
 |||
 Db 976 TGCCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 1035
 QY 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873
 ||
 Db 1036 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 1068

RESULT 8

AAV17816

ID AAV17816 standard; cDNA; 2268 BP.

XX

AC AAV17816;

XX

DT 17-AUG-1998 (first entry)

XX

DE Homo sapiens don-1 gene splice variant.

XX

KW Murine; don-1 gene; melanoma; treatment; adenocarcinoma;

KW epithelial cell; proliferation; stimulation; treatment; tumours;

KW skin; oesophagus; lung; breast; liver; pancreas; colon; prostate;

KW gastrointestinal tract; uterus; wound healing; transmembrane; ss.

XX

OS Homo sapiens.

```

XX
FH   Key                      Location/Qualifiers
FT   CDS                      69..2012
FT                      /*tag=  a
FT                      /note= "don-1 polypeptide"
XX
PN   WO9807736-A1.
XX
PD   26-FEB-1998.
XX
PF   18-AUG-1997;    97WO-US14585.
XX
PR   19-NOV-1996;    96US-0753007.
PR   19-AUG-1996;    96US-0699591.
XX
PA   (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
PI   Busfield SJ,  Gearing DP;
XX
DR   WPI; 1998-169084/15.
DR   P-PSDB; AAW48383.
XX
PT   Mouse and human don-1 polypeptide(s) - useful for treatment of
PT   melanomas and adenocarcinoma(s), and for wound healing
XX
PS   Claim 4; Fig 7; 121pp; English.
XX
CC   The sequence is that of a human don-1 gene splice variant.
CC   Don-1 polypeptides stimulate proliferation of epithelial cells
CC   and thus are implicated in melanomas and adenocarcinomas in which
CC   epithelial cells proliferate out of control. Compounds that
CC   interfere with don-1 mediated cell proliferation can be used
CC   in the treatment of tumours such as melanomas and adenocarcinomas
CC   of the skin, oesophagus, lung, breast, liver, pancreas,
CC   gastrointestinal tract, colon, prostate or uterus. Alternatively,
CC   don-1 polypeptides can be used to stimulate epithelial cell
CC   proliferation, e.g. for wound healing.
XX
SQ   Sequence 2268 BP; 502 A; 735 C; 700 G; 331 T; 0 other;

Query Match          47.7%;  Score 427.8;  DB 19;  Length 2268;
Best Local Similarity  89.8%;  Pred. No. 5.4e-90;
Matches 459;  Conservative  0;  Mismatches  52;  Indels  0;  Gaps  0;

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
      || || | || | | | | || | || | || | | ||
Db      98  CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
      |||||
Db      158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
      |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602

```

```

Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 337
Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
Db      338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397
Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
Db      398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457
Qy      723 CACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
Db      458 CACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517
Qy      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842
Db      518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577
Qy      843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873
Db      578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608

```

RESULT 9

ABS56036

ID ABS56036 standard; cDNA; 1474 BP.

XX

AC ABS56036;

XX

DT 14-JAN-2003 (first entry)

XX

DE cDNA encoding human second splice variant of Don-1.

XX

KW Human; Don-1; epidermal growth factor; EGF; neuregulin;

KW glycoprotein ligand; cell proliferation; cell proliferative disorder;

KW carcinoma; adenocarcinoma cell; myeloma; cell differentiation;

KW cell survival; epithelial cell; wound healing; tumour formation;

KW brain; vulnerary; cytostatic; gene therapy; gene; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 68..1473

FT /*tag= a

FT /partial

FT /product= "Second splice variant of Don-1"

FT /note= "This sequence lacks a stop codon"

FT /transl_except= (pos:107..108, aa:Lys)

FT /note= "This codon has an apparent 1 nucleotide
deletion which alters the reading frame"

XX

PN US2002127594-A1.

XX

PD 12-SEP-2002.

XX

PF 12-MAR-2002; 2002US-0096241.

QY 628 AAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACC 687
 |||
 Db 361 AAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACC 420
 QY 688 GTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCATCCTGGTCGGGGCAC 747
 |||
 Db 421 GTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCATCCTGGTCGGGGCAC 480
 QY 748 GCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTAC 807
 |||
 Db 481 GCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTAC 540
 QY 808 ATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGT 867
 |||
 Db 541 ATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGT 600
 QY 868 CAGCAG 873
 ||
 Db 601 TTGGAG 606

RESULT 10

ABS56045

ID ABS56045 standard; cDNA; 2266 BP.

XX

AC ABS56045;

XX

DT 14-JAN-2003 (first entry)

XX

DE cDNA encoding human third splice variant of Don-1.

XX

KW Human; Don-1; epidermal growth factor; EGF; neuregulin;

KW glycoprotein ligand; cell proliferation; cell proliferative disorder;

KW carcinoma; adenocarcinoma cell; myeloma; cell differentiation;

KW cell survival; epithelial cell; wound healing; tumour formation;

KW brain; vulnerary; cytostatic; gene therapy; gene; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS

68..2010

FT

/*tag= a

FT

/product= "Third splice variant of Don-1"

FT

/transl_except= (pos:107..108, aa:Lys)

FT

/note= "This codon has an apparent 1 nucleotide

FT

deletion which alters the reading frame"

FT

/transl_except= (pos:994..996, aa:Thr)

XX

PN US2002127594-A1.

XX

PD 12-SEP-2002.

XX

PF 12-MAR-2002; 2002US-0096241.

XX

PR 22-JUN-2000; 2000US-0599789.

XX

[illegible]

Db 361 AAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGGAAGGACACC 420

QY 688 GTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCATCCTGGTCGGGGGCAC 747
 |||

Db 421 GTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCATCCTGCTCGGGGCAC 480

QY 748 GCCCGBAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTAC 807
 |||

Db 481 GCCCGBAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTAC 540

QY 808 ATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGT 867
 |||

Db 541 ATCGAGGGCATCAACCAGCTCTCCTGCAATGTCCAATGGATTCTTCGGACAGAGATGT 600

QY 868 CAGCAG 873
 ||

Db 601 TTGGAG 606

RESULT 11

AAV17815

ID AAV17815 standard; cDNA; 1476 BP.

XX

AC AAV17815;

XX

DT 17-AUG-1998 (first entry)

XX

DE Homo sapiens don-1 gene splice variant.

XX

KW Murine; don-1 gene; melanoma; treatment; adenocarcinoma;
 KW epithelial cell; proliferation; stimulation; treatment; tumours;
 KW skin; oesophagus; lung; breast; liver; pancreas; colon; prostate;
 KW gastrointestinal tract; uterus; wound healing; transmembrane; ss.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
FT	CDS	69..1475
FT		/*tag= a
FT		/note= "don-1 polypeptide"

XX

PN WO9807736-A1.

XX

PD 26-FEB-1998.

XX

PF 18-AUG-1997; 97WO-US14585.

XX

PR 19-NOV-1996; 96US-0753007.

PR 19-AUG-1996; 96US-0699591.

XX

PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.

XX

PI Busfield SJ, Gearing DP;

XX

DR WPI; 1998-169084/15.

DR P-PSDB; AAW48382.

XX

PT Mouse and human don-1 polypeptide(s) - useful for treatment of
PT melanomas and adenocarcinoma(s), and for wound healing

XX

PS Claim 4; Fig 4; 121pp; English.

XX

CC The sequence is that of a human don-1 gene splice variant.
CC Don-1 polypeptides stimulate proliferation of epithelial cells
CC and thus are implicated in melanomas and adenocarcinomas in which
CC epithelial cells proliferate out of control. Compounds that
CC interfere with don-1 mediated cell proliferation can be used
CC in the treatment of tumours such as melanomas and adenocarcinomas
CC of the skin, oesophagus, lung, breast, liver, pancreas,
CC gastrointestinal tract, colon, prostate or uterus. Alternatively,
CC don-1 polypeptides can be used to stimulate epithelial cell
CC proliferation, e.g. for wound healing.

XX

SQ Sequence 1476 BP; 335 A; 475 C; 450 G; 216 T; 0 other;

Query Match 47.5%; Score 426.2; DB 19; Length 1476;
Best Local Similarity 89.6%; Pred. No. 1.1e-89;
Matches 458; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY	363	CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG	422
Db	98	CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA	157
QY	423	CGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	482
Db	158	AGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	217
QY	483	ATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	542
Db	218	ATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	277
QY	543	TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA	602
Db	278	TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA	337
QY	603	CTCAGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	662
Db	338	CTCAGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	397
QY	663	CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	722
Db	398	CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	457
QY	723	CACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG	782
Db	458	CACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG	517
QY	783	CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC	842
Db	518	CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAATGTCC	577
QY	843	TGTGGGATACACCGGGGACAGGTGTCAGCAG	873
Db	578	AAATGGATTCTTCGCACAGAGATGTTTGGAG	608

RESULT 12

ABL40993

ID ABL40993 standard; DNA; 1054 BP.

XX

AC ABL40993;

XX

DT 03-JUL-2002 (first entry)

XX

DE Human neuregulin 2 gene exon 1.

XX

KW Human; MUC1; mucin; glycoprotein; cytostatic; cancer; tumour; ECD;
KW extracellular domain; neuregulin 2; isoform; gene; ds.

XX

OS Homo sapiens.

XX

PN WO200222685-A2.

XX

PD 21-MAR-2002.

XX

PF 11-SEP-2001; 2001WO-US28548.

XX

PR 11-SEP-2000; 2000US-231841P.

XX

PA (KUFE/) KUFE D W.

PA (OHNO/) OHNO T.

XX

PI Kufe DW, Ohno T;

XX

DR WPI; 2002-339864/37.

XX

PT Use of a mucin glycoprotein (MUC1) extracellular domain antagonist for
PT manufacturing a medicant that inhibits the proliferation of MUC-1
PT expressing cancer cells and that can treat cancers and reduce tumor
PT growth -

XX

PS Disclosure; Page 61-62; 74pp; English.

XX

CC The invention relates to the use of a MUC1 (mucin glycoprotein)
CC extracellular domain (ECD) antagonist for the manufacture of a medicant
CC to inhibit the proliferation of MUC-1 expressing cancer cells. MUC1 ECD
CC antagonists (optionally combined with a pharmaceutical carrier) can be
CC administered to inhibit proliferation of MUC1-expressing cancer cells,
CC useful to treat cancers e.g. skin cancer, prostate cancer and leukemia,
CC especially in humans. The method may also be combined with administration
CC of a chemotherapeutic agent (e.g. an alkylating agent, topoisomerase etc)
CC or radiation to treat cancer, especially to reduce tumour growth. The
CC polypeptides are also useful in screening to identify MUC1 ECD
CC antagonists. The present sequence represents an exon fragment of the
CC human neuregulin 2 gene.

XX

SQ Sequence 1054 BP; 178 A; 367 C; 361 G; 148 T; 0 other;

Query Match 47.3%; Score 424; DB 24; Length 1054;

Best Local Similarity 100.0%; Pred. No. 3.4e-89;

Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTGC 60
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 589 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTGC 648
 QY 61 TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 649 TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 708
 QY 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 709 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 768
 QY 181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 769 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 828
 QY 241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 829 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 888
 QY 301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 889 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 948
 QY 361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 949 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 1008
 QY 421 TGCG 424
 ||||
 Db 1009 TGCG 1012

RESULT 13

AAV17813

ID AAV17813 standard; cDNA; 1607 BP.

XX

AC AAV17813;

XX

DT 17-AUG-1998 (first entry)

XX

DE Mus musculus don-1 gene splice variant.

XX

KW Murine; don-1 gene; melanoma; treatment; adenocarcinoma;
 KW epithelial cell; proliferation; stimulation; treatment; tumours;
 KW skin; oesophagus; lung; breast; liver; pancreas; colon; prostate;
 KW gastrointestinal tract; uterus; wound healing; secreted protein; ss.

XX

OS Mus musculus.

XX

FH Key Location/Qualifiers
 FT CDS 79..624

FT

FT

XX

PN WO9807736-A1.
 /*tag= a
 /note= "secreted don-1 polypeptide"

```

XX
PD      26-FEB-1998.
XX
PF      18-AUG-1997;   97WO-US14585.
XX
PR      19-NOV-1996;   96US-0753007.
PR      19-AUG-1996;   96US-0699591.
XX
PA      (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
PI      Busfield SJ,   Gearing DP;
XX
DR      WPI; 1998-169084/15.
DR      P-PSDB; AAW48380.
XX
PT      Mouse and human don-1 polypeptide(s) - useful for treatment of
PT      melanomas and adenocarcinoma(s), and for wound healing
XX
PS      Claim 4; Fig 2; 12lpp; English.
XX
CC      The sequence is that of a murine don-1 gene splice variant.
CC      Don-1 polypeptides stimulate proliferation of epithelial cells
CC      and thus are implicated in melanomas and adenocarcinomas in which
CC      epithelial cells proliferate out of control. Compounds that
CC      interfere with don-1 mediated cell proliferation can be used
CC      in the treatment of tumours such as melanomas and adenocarcinomas
CC      of the skin, oesophagus, lung, breast, liver, pancreas,
CC      gastrointestinal tract, colon, prostate or uterus. Alternatively,
CC      don-1 polypeptides can be used to stimulate epithelial cell
CC      proliferation, e.g. for wound healing.
XX
SQ      Sequence 1607 BP; 365 A; 500 C; 480 G; 262 T; 0 other;

Query Match          45.2%; Score 405.4; DB 19; Length 1607;
Best Local Similarity 87.9%; Pred. No. 8.2e-85;
Matches 442; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY      371 CCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 430
          | |||||
Db      2   CTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 61

QY      431 GGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGA 490
          |||||
Db      62 GGCCCAAGCTGAAGAAGATGAAGAGCCAGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCA 121

QY      491 AGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCTACCGTTGGTTCAAGGATGGCAAGG 550
          |||||
Db      122 AGTGTGAGGCAGCGGCGGGAAACCCAGCCCTCTATCGCTGGTTCAAGGATGGCAAGG 181

QY      551 AGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGAC 610
          | |||||
Db      182 AACTCAACCGGAGTCGTGATATTCGCATCAAGTATGGCAATGTCAGAAAGAACTCACGGC 241

QY      611 TACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACA 670
          |||||
Db      242 TACAGTTCAACAAAGTGAGGGTGGAGGATGCCGGGGAGTACGTCTGTGAGGCCGAGAACA 301

```

QY 671 TCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGT 730
 |||| |||||||||||||| |||||| || | |||||||||||||||||| ||||
 Db 302 TCCTTGGGAAGGACACCGTGAGGGGCCGACTCCATGTCAACAGCGTGAGCACCCTCTGT 361
 QY 731 CATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATG 790
 |||||||||||| || |||||||||||||| |||| |||||||||||| || || ||||
 Db 362 CATCCTGGTCGGGACATGCCCGGAAGTGCAATGAGACCGCCAAGTCCTACTGTGTGAATG 421
 QY 791 GAGGCGTCTGCTACTACATCGAGGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGAT 850
 |||||| |||||||||||||||||||||||||||||| |||| ||||
 Db 422 GAGGCGTGTGCTACTACATCGAGGGGCATCAACCAGCTCTCCTGCAAATGTCCAAACGGAT 481
 QY 851 ACACCGGGGACAGGTGTCAGCAG 873
 | ||| | || || ||
 Db 482 TCTTCGGACAGAGATGTTTGGAG 504

RESULT 14

AAV17812

ID AAV17812 standard; cDNA; 2467 BP.

XX

AC AAV17812;

XX

DT 17-AUG-1998 (first entry)

XX

DE Mus musculus don-1 gene splice variant.

XX

KW Murine; don-1 gene; melanoma; treatment; adenocarcinoma;
 KW epithelial cell; proliferation; stimulation; treatment; tumours;
 KW skin; oesophagus; lung; breast; liver; pancreas; colon; prostate;
 KW gastrointestinal tract; uterus; wound healing; transmembrane; ss.

XX

OS Mus musculus.

XX

FH Key Location/Qualifiers

FT CDS

79..1896

FT

/*tag= a

FT

/note= "transmembrane don-1 polypeptide"

XX

PN WO9807736-A1.

XX

PD 26-FEB-1998..

XX

PF 18-AUG-1997; 97WO-US14585.

XX

PR 19-NOV-1996; 96US-0753007.

PR

19-AUG-1996; 96US-0699591.

XX

PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.

XX

PI Busfield SJ, Gearing DP;

XX

DR WPI; 1998-169084/15.

DR

P-PSDB; AAW48379.

XX

PT Mouse and human don-1 polypeptide(s) - useful for treatment of
 PT melanomas and adenocarcinoma(s), and for wound healing

PS Claim 4; Fig 1; 121pp; English.

CC The sequence is that of a murine don-1 gene splice variant.
CC Don-1 polypeptides stimulate proliferation of epithelial cells
CC and thus are implicated in melanomas and adenocarcinomas in which
CC epithelial cells proliferate out of control. Compounds that
CC interfere with don-1 mediated cell proliferation can be used
CC in the treatment of tumours such as melanomas and adenocarcinomas
CC of the skin, oesophagus, lung, breast, liver, pancreas,
CC gastrointestinal tract, colon, prostate or uterus. Alternatively,
CC don-1 polypeptides can be used to stimulate epithelial cell
CC proliferation, e.g. for wound healing.

SQ Sequence 2467 BP; 592 A; 752 C; 706 G; 417 T; 0 other;

Query Match 44.8%; Score 402.2; DB 19; Length 2467;
Best Local Similarity 87.5%; Pred. No. 5.1e-84;
Matches 440; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY	371	CCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC	430
Db	2	CTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC	61
QY	431	GGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGA	490
Db	62	GGCCCAAGCTGAAGAAGATGAAGAGCCAGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCA	121
QY	491	AGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGG	550
Db	122	AGTGTGAGGCAGCGGCGGGAACCCCCAGCCCTCCTATCGCTGGTTCAAGGATGGCAAGG	181
QY	551	AGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGAC	610
Db	182	AACTCAACCGGAGTCGTGATATTTCGCATCAAGTATGGCAATGTCAGAAAGAACTCACGGC	241
QY	611	TACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACA	670
Db	242	TACAGTTCAACAAAGTGAGCGTGAGGATGCCGGGGAGTACGTCTGTGAGGCCGAGAACA	301
QY	671	TCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGT	730
Db	302	TCCTTGGGAAGGACACCGTGAGGGGCCGACTCCATGTCAACAGCGTGACCACCACTCTGT	361
QY	731	CATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATG	790
Db	362	CATCCTGGTCGGGACATGCCCGGAAGTGCAATGAGACCGCCAAGTCCTACTGTGTGAATG	421
QY	791	GAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGAT	850
Db	422	GAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAACGGAT	481
QY	851	ACACCGGGGACAGGTGTCTAGCAG	873
Db	482	TCTTCGGACAGAGATGTTTGGAG	504

RESULT 15

ABS56034

ID ABS56034 standard; cDNA; 1561 BP.

XX

AC ABS56034;

XX

DT 14-JAN-2003 (first entry)

XX

DE cDNA encoding murine secreted splice variant of Don-1.

XX

KW Murine; Don-1; epidermal growth factor; EGF; neuregulin; mouse;
KW glycoprotein ligand; cell proliferation; cell proliferative disorder;
KW carcinoma; adenocarcinoma cell; myeloma; cell differentiation;
KW cell survival; epithelial cell; wound healing; tumour formation;
KW brain; vulnerary; cytostatic; gene therapy; chromosome 18; gene; ss.

XX

OS Mus sp.

XX

FH Key Location/Qualifiers

FT CDS

78..623

FT

/*tag= a

FT

/product= "Secreted splice variant of Don-1"

XX

PN US2002127594-A1.

XX

PD 12-SEP-2002.

XX

PF 12-MAR-2002; 2002US-0096241.

XX

PR 22-JUN-2000; 2000US-0599789.

XX

PA (GEAR/) GEARING D P.

PA

XX

PI Gearing DP, Busfield SJ;

XX

DR WPI; 2003-039584/03.

DR

XX

PT Novel Don-1 polypeptide useful for stimulating proliferation of cells,
PT for identifying proteins that interact with Don-1, and for regulating
PT tumour formation and progression in brain -

XX

PS Claim 4; Fig 2; 66pp; English.

XX

CC The present invention relates to the isolation of a novel gene
CC called Don-1, and alternate splice variants of Don-1, which are
CC related to epidermal growth factors (EGF) such as neuregulins.
CC Don-1 polypeptides are glycoprotein ligands. Both murine and human
CC Don-1 sequences are cloned. The mouse Don-1 gene maps to chromosome 18.
CC Don-1 polypeptides are useful for stimulating proliferation of a cell.
CC Antibodies to Don-1 polypeptides are useful for detecting Don-1
CC in a sample. The Don-1 polypeptides are useful for treating and
CC diagnosing cell proliferative disorders and play a role in the
CC proliferation of carcinomas e.g. adenocarcinoma, myeloma, in cell
CC differentiation, proliferation and survival. The polypeptides are
CC also useful for inhibiting proliferation of adenocarcinoma cells,

CC for stimulating the proliferation of cells such as epithelial cells
CC to promote wound healing, for identifying proteins that interact
CC with Don-1, and for regulating tumour formation and progression in
CC the brain. The polynucleotide sequences encoding Don-1 may be used
CC in gene therapy. The present sequence encodes murine secreted
CC splice variant of Don-1.
XX

SQ Sequence 1561 BP; 361 A; 479 C; 465 G; 256 T; 0 other;

Query Match 43.9%; Score 393.4; DB 25; Length 1561;
Best Local Similarity 87.7%; Pred. No. 5.1e-82;
Matches 441; Conservative 0; Mismatches 61; Indels 1; Gaps 1;

```
QY 371 CCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 430
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2 CTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCA-CC 60
QY 431 GGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGA 490
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 GGCCCAAGCTGAAGAAGATGAAGAGCCAGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCA 120
QY 491 AGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGG 550
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 AGTGTGAGGCAGCGCGGGAAACCCCGCCCTCCTATCGCTGGTTCAAGGATGGCAAGG 180
QY 551 AGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAACTCAGCAG 610
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 AACTCAACCGGAGTCGTGATATTTCGCATCAAGTATGGCAATGTCAGAAAGAACTCAGGCG 240
QY 611 TACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACA 670
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 TACAGTTCAACAAAGTGAGGGTGGAGGATGCCGGGGAGTACGTCTGTGAGGCCGAGAACA 300
QY 671 TCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGT 730
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 TCCTTGGGAAGGACACCGTGAGGGGGCCGACTCCATGTCAACAGCGTGAGCACCCTCTGT 360
QY 731 CATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATG 790
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 CATCCTGGTCGGGACATGCCCGGAAGTGCAATGAGACCGCCAAGTCCTACTGTGTGAATG 420
QY 791 GAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGAT 850
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 GAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAATGTCCAAACGGAT 480
QY 851 ACACCGGGGACAGGTGTCAGCAG 873
    | | | | | | | | | | | | | | | | | |
Db 481 TCTTCGGACAGAGATGTTTGGAG 503
```

Search completed: January 14, 2004, 08:16:39
Job time : 292.304 secs

OM nucleic - nucleic search, using sw model

Run on: January 14, 2004, 07:16:01 ; Search time 66.4093 Seconds
(without alignments)
5961.825 Million cell updates/sec

Title: US-09-864-675-3
Perfect score: 897
Sequence: 1 atgaggcgcgacccggcccc.....caatgggtcaacttctcctaa 897

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	% Query		Length	DB	ID	Description
	Score	Match				
1	835.4	93.1	1884	3	US-08-753-007A-5	Sequence 5, Appli
2	835.4	93.1	1884	3	US-09-398-496-5	Sequence 5, Appli
3	784	87.4	3441	2	US-08-525-864A-1	Sequence 1, Appli
4	738.6	82.3	993	2	US-08-525-864A-3	Sequence 3, Appli
5	427.8	47.7	1476	3	US-08-753-007A-7	Sequence 7, Appli
6	427.8	47.7	1476	3	US-09-398-496-7	Sequence 7, Appli
7	427.8	47.7	2268	3	US-08-753-007A-31	Sequence 31, Appl
8	427.8	47.7	2268	3	US-09-398-496-31	Sequence 31, Appl
9	405.4	45.2	1607	3	US-08-753-007A-3	Sequence 3, Appli
10	405.4	45.2	1607	3	US-09-398-496-3	Sequence 3, Appli
11	405.4	45.2	2467	3	US-08-753-007A-1	Sequence 1, Appli

12	405.4	45.2	2467	3	US-09-398-496-1	Sequence 1, Appli
13	216.2	24.1	1207	2	US-08-525-864A-5	Sequence 5, Appli
14	84	9.4	2003	1	US-08-036-555B-21	Sequence 21, Appl
15	84	9.4	2003	1	US-08-469-569-21	Sequence 21, Appl
16	84	9.4	2003	1	US-08-249-322A-21	Sequence 21, Appl
17	84	9.4	2003	1	US-08-469-526A-21	Sequence 21, Appl
18	84	9.4	2003	2	US-08-734-591A-21	Sequence 21, Appl
19	84	9.4	2003	2	US-08-469-660-21	Sequence 21, Appl
20	84	9.4	2003	3	US-08-341-018-71	Sequence 71, Appl
21	84	9.4	2003	3	US-08-470-335-21	Sequence 21, Appl
22	84	9.4	2003	3	US-08-735-021-21	Sequence 21, Appl
23	84	9.4	2003	3	US-08-734-664A-21	Sequence 21, Appl
24	84	9.4	2003	3	US-08-470-339-21	Sequence 21, Appl
25	84	9.4	2003	4	US-08-467-602-21	Sequence 21, Appl
26	84	9.4	2003	5	PCT-US94-05083C-21	Sequence 21, Appl
27	84	9.4	2003	5	PCT-US95-06846A-21	Sequence 21, Appl
28	83.4	9.3	1108	1	US-08-036-555B-135	Sequence 135, App
29	83.4	9.3	1108	1	US-08-469-569-135	Sequence 135, App
30	83.4	9.3	1108	1	US-08-249-322A-135	Sequence 135, App
31	83.4	9.3	1108	1	US-08-469-526A-135	Sequence 135, App
32	83.4	9.3	1108	2	US-08-734-591A-135	Sequence 135, App
33	83.4	9.3	1108	2	US-08-469-660-135	Sequence 135, App
34	83.4	9.3	1108	3	US-08-341-018-5	Sequence 5, Appli
35	83.4	9.3	1108	3	US-08-470-335-135	Sequence 135, App
36	83.4	9.3	1108	3	US-08-735-021-135	Sequence 135, App
37	83.4	9.3	1108	3	US-08-734-664A-135	Sequence 135, App
38	83.4	9.3	1108	3	US-08-470-339-135	Sequence 135, App
39	83.4	9.3	1108	4	US-08-467-602-135	Sequence 135, App
40	83.4	9.3	1108	5	PCT-US94-05083C-131	Sequence 131, App
41	83.4	9.3	1108	5	PCT-US95-06846A-135	Sequence 135, App
42	78.4	8.7	1193	1	US-08-469-526A-134	Sequence 134, App
43	78.4	8.7	1193	2	US-08-734-591A-134	Sequence 134, App
44	78.4	8.7	1193	3	US-08-341-018-3	Sequence 3, Appli
45	78.4	8.7	1193	3	US-08-470-335-134	Sequence 134, App

ALIGNMENTS

RESULT 1

US-08-753-007A-5

; Sequence 5, Application US/08753007A

; Patent No. 6074841

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/753,007A
; FILING DATE: 19-NOV-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1884 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 664...1883
; OTHER INFORMATION:
US-08-753-007A-5

```

```

Query Match          93.1%; Score 835.4; DB 3; Length 1884;
Best Local Similarity 98.1%; Pred. No. 8.6e-200;
Matches 856; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

```

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      218 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 277
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTTGGTGGTGGAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTTGGTGGTGGAG 337
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      338 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 397
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy      181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      398 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 457
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy      241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      458 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 517
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy      301 CGCTACATCTTTTCTGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360

```

Db	518	CGCTACATCTTTTCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	577
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	578	CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC	636
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	637	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	696
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	697	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	756
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	757	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG	816
Qy	601	AACTCAGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	817	AACTCAGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	876
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	877	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	936
Qy	721	ACCACCCTGTCATCCTGGTTCGGGGCAGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	937	ACCACCCTGTCATCCTGGTTCGGGGCAGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	996
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT	840
Db	997	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	1056
Qy	841	CCTGTGGGATACACCGGGACAGGTGTCAGCAG	873
Db	1057	CCAAATGGATTCTTCGGACAGAGATGTTTGAG	1089

RESULT 2

US-09-398-496-5

; Sequence 5, Application US/09398496

; Patent No. 6133423

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/398,496
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/753,007
; FILING DATE: 19-NOV-1996
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1884 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 664...1883
; OTHER INFORMATION:
US-09-398-496-5

```

```

Query Match          93.1%; Score 835.4; DB 3; Length 1884;
Best Local Similarity 98.1%; Pred. No. 8.6e-200;
Matches 856; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

```

```

QY      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 60
        |||
Db      218 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 277
        |||
QY      61 TACTCGCCCGACCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        |||
Db      278 TACTCGCCCGACCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 337
        |||
QY      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
        |||
Db      338 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 397
        |||
QY      181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
        |||
Db      398 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 457
        |||
QY      241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
        |||

```

Db 458 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 517
 QY 301 CGCTACATCTTTTTCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 518 CGCTACATCTTTTTCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 577
 QY 361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
 ||||| |||||||||||||||||||||||||||||||||||||||||||||
 Db 578 CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC 636
 QY 421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 637 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 696
 QY 481 CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 697 CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 756
 QY 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 600
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 757 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 816
 QY 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 817 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 876
 QY 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 877 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 936
 QY 721 ACCACCCTGTCATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 937 ACCACCCTGTCATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 996
 QY 781 TCGGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 997 TCGGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 1056
 QY 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873
 || ||||| ||| ||| ||| |||
 Db 1057 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 1089

RESULT 3

US-08-525-864A-1

; Sequence 1, Application US/08525864A

; Patent No. 5912326

; GENERAL INFORMATION:

; APPLICANT: Chang, Han

; TITLE OF INVENTION: Cerebellum-derived Growth Factors, and Uses

; TITLE OF INVENTION: Related thereto

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts


```

; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: AscII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,864A
; FILING DATE: 8-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kara, Catherine J.
; REGISTRATION NUMBER: 41,106
; REFERENCE/DOCKET NUMBER: HUI-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3441 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 180..2441
US-08-525-864A-1

```

```

Query Match          87.4%; Score 784; DB 2; Length 3441;
Best Local Similarity 92.2%; Pred. No. 7.5e-187;
Matches 826; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```

```

QY      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      180 ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 239
        |||||||||||||||||||||||||||||||||||||||||||||||||||
QY      61 TACTCGCCAGCCTCAAGTCAGTGAGGACAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      240 TACTCGCCAGCCTCAAGTCCGTGCAGGACAGGCGTACAAGGCACCCGTGGTGGTGGAG 299
        |||||||||||||||||||||||||||||||||||||||||||||||||||
QY      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
        |||||||||||||| |||| ||| || |||| |||||||| |||||||||||
Db      300 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCAGCTCTAACAGCACCCGAGAGCCT 359
        |||||||||||||||||||||||||||||||||||||||||||||||||||
QY      181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
        |||||||||||||||||| |||| |||||||||||||||||||||||||||
Db      360 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 419
        |||||||||||||||||||||||||||||||||||||||||||||||||||
QY      241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      420 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 479
        |||||||||||||||||||||||||||||||||||||||||||||||||||
QY      301 CGCTACATCTTTTTCTTGGAGCCACCGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
        |||||||||||||||||||||| || |||||||||| |||||||||||
Db      480 CGCTACATCTTTTTCTTGGAGCCACCGAGCAGCCCTTAGTTTTTAAGACAGCCTTTGCC 539

```



```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: AscII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,864A
; FILING DATE: 8-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kara, Catherine J.
; REGISTRATION NUMBER: 41,106
; REFERENCE/DOCKET NUMBER: HUI-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..990
US-08-525-864A-3

```

```

Query Match      82.3%; Score 738.6; DB 2; Length 993;
Best Local Similarity 90.4%; Pred. No. 1.1e-175;
Matches 789; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

```

```

QY      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 60
Db      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 60
QY     61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
Db     61 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
QY    121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
Db    121 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT 180
QY    181 CCCGCCTCGGGTTCGGGTGGCGTTGCTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
Db    181 CCCGCCTCGGGTTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
QY    241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
Db    241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 300
QY    301 CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
Db    301 CGCTACATCTTTTTCCTGGAGCCCACCGAGCAGCCCTTAGTCTTTAAGACAGCCTTTGCC 360
QY    361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
Db    361 CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420

```


FILING DATE: 19-NOV-1996
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/699,591
 FILING DATE: 19-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Fasse, J. Peter
 REGISTRATION NUMBER: 32,983
 REFERENCE/DOCKET NUMBER: 07334/022001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-542-5070
 TELEFAX: 617-542-8906
 TELEX:
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1476 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 69...1475
 OTHER INFORMATION:
 US-08-753-007A-7

Query Match 47.7%; Score 427.8; DB 3; Length 1476;
 Best Local Similarity 89.8%; Pred. No. 6.9e-98;
 Matches 459; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy	363	CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG	422
Db			
	98	CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA	157
Qy	423	CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	482
Db			
	158	AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	217
Qy	483	ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	542
Db			
	218	ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	277
Qy	543	TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA	602
Db			
	278	TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA	337
Qy	603	CTCAGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	662
Db			
	338	CTCAGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	397
Qy	663	CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	722
Db			
	398	CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	457
Qy	723	CACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG	782
Db			
	458	CACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG	517

QY 783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGGCATCAACCAGCTCTCCTGCAAGTGTC 842
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGGCATCAACCAGCTCTCCTGCAAATGTCC 577
 QY 843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608

RESULT 6

US-09-398-496-7

; Sequence 7, Application US/09398496
 ; Patent No. 6133423

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/398,496

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/753,007

; FILING DATE: 19-NOV-1996

; APPLICATION NUMBER: 08/699,591

; FILING DATE: 19-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Fasse, J. Peter

; REGISTRATION NUMBER: 32,983

; REFERENCE/DOCKET NUMBER: 07334/022001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-542-5070

; TELEFAX: 617-542-8906

; TELEX:

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1476 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

```

;      NAME/KEY:   Coding Sequence
;      LOCATION:   69...1475
;      OTHER INFORMATION:
US-09-398-496-7

```

Query Match

Best Local Similarity 47.7%; Score 427.8; DB 3; Length 1476;
Matches 459; Conservative 89.8%; Pred. No. 6.9e-98;

Similarity 89.8%; Pred. No. 6.9e-98; Length 1476;
Matches 459; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
363 CCTCGATACCAACGGCAAAATTC

Qy 363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
 Db ||| || | || | || | ||| ||| ||| ||| |||

Db
98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 422
QY
423 CGCCACCCGGCCCAAGTTGAAGAGGCTTCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 157

Qy 423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 157
Db 158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482

Db
158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
Qy
483 ATCGCTGAAGTGTGAGGCAGCCAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

QY 483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 217
Db 218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542

Db
218 ATCGCTGAAGTGTGAGGCAGCAGCCGGAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
QY
543 TGGCAAGGAGCTCAACCGCAGCCGGAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy 543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 277
Db 278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602

Db
278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602
Qy
603 CTCACGACTACAGTTCAACAAGCTG 337

QY 603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 337
Db 338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662

Db
338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
Qy
663 CGAGAACATCCTGGGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

QY 663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 397
Db 398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722

Db
398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
QY
723 CACCCTGTGATCCTGCTGAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

QY 723 CACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 457

Db 458 CACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782

Dd
458 CACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
QY
783 CGTCAATGGAGGCCTCTGCTTTT

QY 783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 517
 Db 518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842

Db
518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842
Qy
843 TGTGGGATACAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577

Qy 843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873
Db 578 AAATGCATGCTTCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577

Db
578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608

RESULT 7

US-08-753-007A-31

; Sequence 31, Application US/08753007A
; Patent No. 6074841

Patent No. 6074841
GENERAL INFORMATIO

GENERAL INFORMATION:
APPLICANT: Gearing, David P.
APPLICANT: Bugfi

APPLICANT: Gearing, David P.
; APPLICANT: Busfield, Samantha J.
; TITLE OF INVENTION: BOX 1

TITLE OF INVENTION: BUSFIELD, Samantha J.
 TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
 NUMBER OF INVENTION: AND USES THEREOF

; TITLE OF INVENTION: DON-1 GENE AND PO
 ; NUMBER OF INVENTION: AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 33
 ; ADDRESS:

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: US
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/753,007A
 FILING DATE: 19-NOV-1996
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/699,591
 FILING DATE: 19-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Fasse, J. Peter
 REGISTRATION NUMBER: 32,983
 REFERENCE/DOCKET NUMBER: 07334/022001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-542-5070
 TELEFAX: 617-542-8906
 TELEX:
 INFORMATION FOR SEQ ID NO: 31:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2268 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 69...2009
 OTHER INFORMATION:
 US-08-753-007A-31

Query Match 47.7%; Score 427.8; DB 3; Length 2268;
 Best Local Similarity 89.8%; Pred. No. 7.9e-98;
 Matches 459; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy	363	CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG	422
Db	98	CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA	157
Qy	423	CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	482
Db	158	AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	217
Qy	483	ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	542
Db	218	ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	277
Qy	543	TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA	602

Db 278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 337
 QY 603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
 Db 338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397
 QY 663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
 Db 398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457
 QY 723 CACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
 Db 458 CACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517
 QY 783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842
 Db 518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577
 QY 843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873
 Db 578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608

RESULT 8

US-09-398-496-31

; Sequence 31, Application US/09398496
 ; Patent No. 6133423

GENERAL INFORMATION:

APPLICANT: Gearing, David P.

APPLICANT: Busfield, Samantha J.

TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

TITLE OF INVENTION: AND USES THEREFOR

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/398,496

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/753,007

FILING DATE: 19-NOV-1996

APPLICATION NUMBER: 08/699,591

FILING DATE: 19-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Fasse, J. Peter

REGISTRATION NUMBER: 32,983

REFERENCE/DOCKET NUMBER: 07334/022001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX:
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 2268 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 69...2009
OTHER INFORMATION:
US-09-398-496-31

Query Match 47.7%; Score 427.8; DB 3; Length 2268;
Best Local Similarity 89.8%; Pred. No. 7.9e-98;
Matches 459; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy	363	CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG	422
Db			
Qy	98	CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA	157
Db			
Qy	423	CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	482
Db			
Qy	158	AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	217
Db			
Qy	483	ATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	542
Db			
Qy	218	ATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	277
Db			
Qy	543	TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA	602
Db			
Qy	278	TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA	337
Db			
Qy	603	CTCAGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	662
Db			
Qy	338	CTCAGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	397
Db			
Qy	663	CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	722
Db			
Qy	398	CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	457
Db			
Qy	723	CACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG	782
Db			
Qy	458	CACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG	517
Db			
Qy	783	CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC	842
Db			
Qy	518	CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAAATGTCC	577
Db			
Qy	843	TGTGGGATACACCGGGGACAGGTGTCAGCAG	873
Db			
	578	AAATGGATTCTTCGGACAGAGATGTTTGGAG	608

RESULT 9

US-08-753-007A-3

; Sequence 3, Application US/08753007A
; Patent No. 6074841

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.
; APPLICANT: Busfield, Samantha J.
; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/753,007A
; FILING DATE: 19-NOV-1996
; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX:

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1607 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: Coding Sequence
; LOCATION: 79...621

; OTHER INFORMATION:

US-08-753-007A-3

Query Match

45.2%; Score 405.4; DB 3; Length 1607;

Best Local Similarity

87.9%; Pred. No. 2.8e-92;

Matches 442; Conservative

0; Mismatches 61; Indels 0; Gaps 0;

QY

371 CCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 430

Db 2 CTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 61
 QY 431 GGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGA 490
 Db 62 GGCCCAAGCTGAAGAAGATGAAGAGCCAGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCA 121
 QY 491 AGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGG 550
 Db 122 AGTGTGAGGCAGCGGGGAAACCCCCAGCCCTCCTATCGCTGGTTCAAGGATGGCAAGG 181
 QY 551 AGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGAC 610
 Db 182 AACTCAACCGGAGTCGTGATATTCGCATCAAGTATGGCAATGTGAGAAAGAACTCACGGC 241
 QY 611 TACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACA 670
 Db 242 TACAGTTCAACAAAGTGAGGGTGGAGGATGCCGGGGAGTACGTCTGTGAGGCCGAGAACA 301
 QY 671 TCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGT 730
 Db 302 TCCTTGGGAAGGACACCGTGAGGGGCCGACTCCATGTCAACAGCGTGAGCACCCTCTGT 361
 QY 731 CATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATG 790
 Db 362 CATCCTGGTCGGGACATGCCCGGAAGTGCAATGAGACCGCCAAGTCCTACTGTGTGAATG 421
 QY 791 GAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGAT 850
 Db 422 GAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAACGGAT 481
 QY 851 ACACCGGGGACAGGTGTCAGCAG 873
 Db 482 TCTTCGGACAGAGATGTTTGGAG 504

RESULT 10

US-09-398-496-3

; Sequence 3, Application US/09398496

; Patent No. 6133423

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

Db 302 TCCTTGGAAGGACACCGTGAGGGGCCGACTCCATGTCAACAGCGTGAGCACCCTCTGT 361

QY 731 CATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATG 790

Db 362 CATCCTGGTCGGGACATGCCCCGGAAGTGCAATGAGACCGCCAAGTCCTACTGTGTGAATG 421

QY 791 GAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGAT 850

Db 422 GAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAACGGAT 481

QY 851 ACACCGGGGACAGGTGTCAGCAG 873

Db 482 TCTTCGGACAGAGATGTTTGGAG 504

RESULT 11

US-08-753-007A-1

```

; Sequence 1, Application US/08753007A
; Patent N. 6854

```

; Patent No. 6074841

; GENERAL INFORMATION:

APPLICANT: Gearing, David P.

APPLICANT: Busfield, Samantha J.

TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR
 ; NUMBER OF SECTIONS:

```

; NUMBER OF SEQUENCES: 33

```

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/753,007A

FILING DATE: 19-NOV-1996

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/699,591

FILING DATE: 19-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Fasse, J. Peter

REGISTRATION NUMBER: 32,983

REFERENCE/DOCKET NUMBER: 07334/022001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-542-5070

TELEFAX: 617-542-8906

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2467 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

US-08-753-007A-1

Matches 442; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

RESULT 12

; Sequence 1, Application US/09398496

; GENERAL INFORMATION:

APPLICANT: Busfield, Samantha J.

TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

Db 122 AGTGTGAGGCAGCGGCGGGAAACCCCGAGCCCTCCTATCGCTGGTTCAAGGATGGCAAGG 181
 Qy 551 AGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGAC 610
 Db 182 AACTCAACCGGAGTCGTGATATTTCGCATCAAGTATGGCAATGTCAGAAAGAACTCACGGC 241
 Qy 611 TACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACA 670
 Db 242 TACAGTTCAACAAAGTGAGGGTGGAGGATGCCGGGGAGTACGTCTGTGAGGCCGAGAACA 301
 Qy 671 TCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGT 730
 Db 302 TCCTTGGGAAGGACACCGTGAGGGGCCGACTCCATGTCAACAGCGTGAGCACCCTGT 361
 Qy 731 CATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCTATTGCGTCAATG 790
 Db 362 CATCCTGGTCGGGACATGCCCCGGAAGTGCAATGAGACCGCCAAGTCTACTGTGTGAATG 421
 Qy 791 GAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGAT 850
 Db 422 GAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAATGTCCAACGGAT 481
 Qy 851 ACACCGGGGACAGGTGTGTCAGCAG 873
 Db 482 TCTTCGGACAGAGATGTTTGGAG 504

RESULT 13

US-08-525-864A-5

; Sequence 5, Application US/08525864A

; Patent No. 5912326

; GENERAL INFORMATION:

; APPLICANT: Chang, Han

; TITLE OF INVENTION: Cerebellum-derived Growth Factors, and Uses

; TITLE OF INVENTION: Related thereto

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: AscII (text)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/525,864A

; FILING DATE: 8-SEP-1995

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Kara, Catherine J.

; REGISTRATION NUMBER: 41,106

; REFERENCE/DOCKET NUMBER: HUI-017

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..394
US-08-525-864A-5

```

```

Query Match          24.1%; Score 216.2; DB 2; Length 1207;
Best Local Similarity 86.3%; Pred. No. 5.4e-45;
Matches 239; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

```

```

QY      597 AAAGAACTCAGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTG 656
      |||
Db      1 AAAGAACTCAGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTG 60
      |||
QY      657 CGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGT 716
      |||
Db      61 TGAGGCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGT 120
      |||
QY      717 GAGCACCACCCTGTCTCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTC 776
      |||
Db      121 GAGCACCCTCTGTCTCCTGGTTCGGGGCAGCCCCGGAAGTGCAATGAGACAGCCAAGTC 180
      |||
QY      777 CTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAA 836
      |||
Db      181 CTACTGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAA 240
      |||
QY      837 GTGTCCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873
      |||
Db      241 ATGTCCAAACGGATTCTTCGGACAGAGATGTTTGGAG 277

```

RESULT 14

US-08-036-555B-21

; Sequence 21, Application US/08036555B

; Patent No. 5530109

; GENERAL INFORMATION:

; APPLICANT: Goodearl, Andrew; Stroobant, Paul;

; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;

; APPLICANT: Chen, Maio Su; Hiles, Ian

; TITLE OF INVENTION: Glial Mitogenic Factors, Their

; TITLE OF INVENTION: Preparation and Use

; NUMBER OF SEQUENCES: 184

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

```

; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/036,555B
; FILING DATE: 24-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 5250.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2003
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: N in positions 31 and 32 could be either
; OTHER INFORMATION: A or G.
US-08-036-555B-21

```

Query Match 9.4%; Score 84; DB 1; Length 2003;

Best Local Similarity 49.0%; Pred. No. 7.3e-12;

Matches 361; Conservative 0; Mismatches 345; Indels 30; Gaps 4;

```

QY      189 GGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGGGGGCTGCA 248
      ||  ||| || |  |  ||  || | |||  |||| | ||| |||
Db      762 GGTGTGGGCGGTGAAAGCCGGGGGCTTGAAGAAGGACTCGCTGCTCACCGTGCCTGGG 821
QY      249 GCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCGCTACAT 308
      | | |  |  |  ||| ||| | ||| |  ||  | ||||
Db      822 GACCTGGGGCCACCCCGCCTTCCCCTCCTGCGGGAGGCTCAAGGAGGACAGCAGGTACAT 881
QY      309 CTTTTTCCTGGAGCCCCACGGAACAGCCCTTAGTCTTTAAGACG-----GCCTT 356
      ||| ||| ||||| ||  |  |  |  ||  |||

```


CITY: New York City
 STATE: New York
 COUNTRY: USA
 ZIP: 10022
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
 COMPUTER: IBM
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: Wordperfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/469,569
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/036,555
 FILING DATE: 24-MAR-1993
 APPLICATION NUMBER: 07/965,173
 FILING DATE: 23-OCT-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/940,389
 FILING DATE: 03-SEP-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/907,138
 FILING DATE: 30-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/863,703
 FILING DATE: 03-APRIL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: U.K. 91 07566.3
 FILING DATE: 10-APRIL-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Tsai, Christine H.
 REGISTRATION NUMBER: 34,266
 REFERENCE/DOCKET NUMBER: LUD 5250.4
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2003
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 OTHER INFORMATION: N in positions 31 and 32 could be either
 OTHER INFORMATION: A or G.
 US-08-469-569-21

Query Match 9.4%; Score 84; DB 1; Length 2003;
 Best Local Similarity 49.0%; Pred. No. 7.3e-12;
 Matches 361; Conservative 0; Mismatches 345; Indels 30; Gaps 4;

QY 189 GGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGCCGCTCCGGAGCGGGGGGCTGCA 248
 || ||| || | | || ||| | | | | |
 Db 762 GGTGTGGGCGGTGAAAGCCGGGGGCTTGAAGAAGGACTCGCTGCTCACC GTGCGCCTGGG 821
 QY 249 GCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAGCGCTACAT 308

Db 822 GACCTGGGGCCACCCCGCCTTCCCTCCTGCGGGAGGCTCAAGGAGGACAGCAGGTACAT 881
 QY 309 CTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACG-----GCCTT 356
 Db 882 CTTCTTCATGGAGCCCGACGCCAACAGCACCAGCCGCGCGCCGCGCCTTCCGAGCCTC 941
 QY 357 TGCCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCAC 416
 Db 942 TTTCCCCCTCTGGAGACGGGCCGGAACCTCAAGAAGGAGGTGAGCCGGGTGCTGTGCAA 1001
 QY 417 TGA CTGCGCCACCCGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGA 476
 Db 1002 GCGGTGCGCCTTGCTCCCCAATTGAAAGAGATGAAAGCCAGGAATCGGCTGCAGGTTT 1061
 QY 477 GAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCTACCGTTGGTT 536
 Db 1062 CAACTAGTCCTTCGGTGTGAAACAGTTCTGAATACTCCTCTCTCAGATTCAAGTGTT 1121
 QY 537 CAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCA---TCAAATATGGCAACGG 593
 Db 1122 CAAGAATGGGAATGAATTGAATCGAAAAACAAACCACAAATATCAAGATACAAAAAA 1181
 QY 594 CAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGT 653
 Db 1182 GCCAGGGAAGTCAGAACTTCGCATTAACAAAGCATCACTGGCTGATTCTGGAGAGTATAT 1241
 QY 654 CTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAG 713
 Db 1242 GTGCAAAGTGATCAGCAAATTAGGAAATGACAGTGCCTCTGCCAATATCACCATCGTGGA 1301
 QY 714 CGTGAGCACCACTTGTCTATCC-----TGGTGGGGGACGCCCCGGAAGTGCAACGAGAC 767
 Db 1302 ATCAAACGCTACATCTACATCCACCCTGGGACAAGCCATCTTGTAATATGTGCGGAGAA 1361
 QY 768 AGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCT 827
 Db 1362 GGAGAAAACCTTCTGTGTGAATGGAGGGGAGTGCTTCATGGTGAAAGACCTTTCAAACCC 1421
 QY 828 CTC-----CTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTTCG 878
 Db 1422 CTCGAGATACTTGTGCAAGTGCCCAATGAGTTTACTGGTGATCGCTGCCAAAACCTACGT 1481
 QY 879 AATGGTCAACTTCTCC 894
 Db 1482 AATGGCCAGCTTCTAC 1497

Search completed: January 14, 2004, 10:25:25
 Job time : 68.4093 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2004, 10:23:13 ; Search time 346.751 Seconds
(without alignments)
9118.095 Million cell updates/sec

Title: US-09-864-675-3
Perfect score: 897
Sequence: 1 atgaggcgcgacccggccccc.....caatgggtcaacttctcctaa 897

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Match	Length	DB	ID	Description
No.						

	1	897	100.0	897	9	US-09-864-675-3	Sequence 3, Appli
	2	849	94.6	994	9	US-09-864-675-1	Sequence 1, Appli
	3	835.4	93.1	1884	14	US-10-096-241-5	Sequence 5, Appli
	4	427.8	47.7	1476	14	US-10-096-241-7	Sequence 7, Appli
	5	427.8	47.7	2268	14	US-10-096-241-31	Sequence 31, Appl
	6	405.4	45.2	1607	14	US-10-096-241-3	Sequence 3, Appli
	7	405.4	45.2	2467	14	US-10-096-241-1	Sequence 1, Appli
	8	173	19.3	201	13	US-10-029-386-26613	Sequence 26613, A
	9	173	19.3	573	13	US-10-029-386-12913	Sequence 12913, A
c	10	113.6	12.7	579	13	US-10-029-386-2532	Sequence 2532, Ap
c	11	111.8	12.5	171	13	US-10-029-386-16232	Sequence 16232, A
	12	84	9.4	1986	11	US-09-373-658-71	Sequence 71, Appl
	13	84	9.4	1986	12	US-09-989-687-71	Sequence 71, Appl
	14	84	9.4	2003	7	US-08-736-019-21	Sequence 21, Appl
	15	84	9.4	2003	11	US-09-366-886-71	Sequence 71, Appl
	16	84	9.4	2003	11	US-09-373-658-72	Sequence 72, Appl
	17	84	9.4	2003	12	US-09-989-687-72	Sequence 72, Appl
	18	83.4	9.3	1108	7	US-08-736-019-135	Sequence 135, App
	19	83.4	9.3	1108	11	US-09-366-886-5	Sequence 5, Appli
	20	78.4	8.7	1193	7	US-08-736-019-134	Sequence 134, App
	21	78.4	8.7	1193	11	US-09-366-886-3	Sequence 3, Appli
	22	76.8	8.6	1140	7	US-08-736-019-149	Sequence 149, App
	23	76.8	8.6	1140	11	US-09-366-886-55	Sequence 55, Appl
	24	72.8	8.1	744	7	US-08-736-019-133	Sequence 133, App
	25	72.8	8.1	744	11	US-09-366-886-1	Sequence 1, Appli
	26	58.8	6.6	848	13	US-10-027-632-158930	Sequence 158930,
	27	58.8	6.6	848	14	US-10-027-632-158930	Sequence 158930,
	28	52	5.8	1092	15	US-10-156-761-4747	Sequence 4747, Ap
	29	52	5.8	9025608	15	US-10-156-761-1	Sequence 1, Appli
	30	51	5.7	162	15	US-10-076-816-8	Sequence 8, Appli
	31	51	5.7	162	15	US-10-077-381-8	Sequence 8, Appli
	32	50.6	5.6	3111	9	US-09-773-517-12	Sequence 12, Appl
	33	50.6	5.6	3111	9	US-09-792-025-12	Sequence 12, Appl
	34	50.6	5.6	3111	9	US-09-849-868-12	Sequence 12, Appl
	35	50.6	5.6	3111	13	US-10-453-183-12	Sequence 12, Appl
	36	50.6	5.6	3111	15	US-10-290-578-1	Sequence 1, Appli
	37	49.8	5.6	1935	10	US-09-778-510-21	Sequence 21, Appl
	38	49.8	5.6	1935	10	US-09-778-187B-3	Sequence 3, Appli
	39	49.8	5.6	1935	13	US-10-302-041-21	Sequence 21, Appl
	40	46.8	5.2	1652	7	US-08-736-019-148	Sequence 148, App
	41	46.8	5.2	1652	11	US-09-366-886-53	Sequence 53, Appl
	42	46.2	5.2	162	12	US-10-428-339-38	Sequence 38, Appl
	43	46	5.1	1521	15	US-10-156-761-5445	Sequence 5445, Ap
	44	45.8	5.1	543	11	US-09-918-995-26620	Sequence 26620, A
	45	45.4	5.1	1092	15	US-10-156-761-6888	Sequence 6888, Ap

ALIGNMENTS

RESULT 1
 US-09-864-675-3
 ; Sequence 3, Application US/09864675
 ; Patent No. US20020081286A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Marchionni, Mark


```

Db      541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 600
Qy      601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
Qy      661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
Qy      721 ACCACCCTGTCATCCTGGTCCGGGGCACGCCCAGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 ACCACCCTGTCATCCTGGTCCGGGGCACGCCCAGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
Qy      781 TCGGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 TCGGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
Qy      841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTCCTAA 897
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTCCTAA 897

```

RESULT 2

US-09-864-675-1

; Sequence 1, Application US/09864675

; Patent No. US20020081286A1

; GENERAL INFORMATION:

; APPLICANT: Marchionni, Mark

; TITLE OF INVENTION: NRG-2 NUCLEIC ACID MOLECULES,

; TITLE OF INVENTION: POLYPEPTIDES, AND DIAGNOSTIC AND THERAPEUTIC METHODS

; FILE REFERENCE: 04585/049002

; CURRENT APPLICATION NUMBER: US/09/864,675

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/206,495

; PRIOR FILING DATE: 2000-05-23

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 994

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-864-675-1

Query Match 94.6%; Score 849; DB 9; Length 994;

Best Local Similarity 98.3%; Pred. No. 1.8e-229;

Matches 858; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60
Qy      61 TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
Qy      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180

```

Db	121		GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Qy	181		CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	181		CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Qy	241		GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	241		GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Qy	301		CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	301		CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Qy	361		CCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	361		CCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Qy	421		TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	421		TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Qy	481		CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	481		CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Qy	541		GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG	600
Db	541		GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG	600
Qy	601		AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	601		AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Qy	661		GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	661		GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Qy	721		ACCACCCTGTCATCCTGGTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTAT	780
Db	721		ACCACCCTGTCATCCTGGTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTAT	780
Qy	781		TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT	840
Db	781		TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT	840
Qy	841		CCTGTGGGATACACCGGGGACAGGTGTCAGCAG	873
Db	841		CCTGTGGGATACACCGGGGACAGGTGTCAGCAG	873

RESULT 3

US-10-096-241-5

; Sequence 5, Application US/10096241

; Publication No. US20020127594A1

; GENERAL INFORMATION:

```

; APPLICANT: Gearing, David P.
; Busfield, Samantha J.
; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
; AND USES THEREFOR
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/096,241
; FILING DATE: 12-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1884 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 664...1883
; OTHER INFORMATION:
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-096-241-5

```

```

Query Match          93.1%; Score 835.4; DB 14; Length 1884;
Best Local Similarity 98.1%; Pred. No. 1.4e-225;
Matches 856; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

```

```

Qy      1 ATGAGGCGCGACCCGGCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      218 ATGAGGCGCGACCCGGCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 277

Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 337

```

Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	338	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	397
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	398	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	457
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	458	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	517
Qy	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	518	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	577
Qy	361	CCCCTCGATACCAACGGCAAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	578	CCCCT-GATACCAACGGCAAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC	636
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	637	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	696
Qy	481	CAATCGCTGAAGTGTGAGGCGAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	697	CAATCGCTGAAGTGTGAGGCGAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	756
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG	600
Db	757	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG	816
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	817	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	876
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	877	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	936
Qy	721	ACCACCCTGTCATCCTGGTCCGGGACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	937	ACCACCCTGTCATCCTGGTCCGGGACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	996
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	840
Db	997	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	1056
Qy	841	CCTGTGGGATACACCGGGACAGGTGTCAGCAG	873
Db	1057	CCAAATGGATTCTTCGGACAGAGATGTTTGGAG	1089

RESULT 4

US-10-096-241-7

; Sequence 7, Application US/10096241

```

; Publication No. US20020127594A1
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; Busfield, Samantha J.
; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
; AND USES THEREFOR
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/096,241
; FILING DATE: 12-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1476 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 69...1475
; OTHER INFORMATION:
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-096-241-7

```

```

Query Match          47.7%; Score 427.8; DB 14; Length 1476;
Best Local Similarity 89.8%; Pred. No. 1.4e-110;
Matches 459; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```

```

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
        || || | || | || | || | || | || | ||
Db      98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482

```

```

      |||
Db      158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
      |||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602
      |||
Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 337

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
      |||
Db      338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
      |||
Db      398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy      723 CACCCTGTCTATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
      |||
Db      458 CACCCTGTCTATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517

Qy      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842
      |||
Db      518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGGCATCAACCAGCTCTCCTGCAATGTCC 577

Qy      843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873
      |||
Db      578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608

```

RESULT 5

US-10-096-241-31

; Sequence 31, Application US/10096241

; Publication No. US20020127594A1

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/096,241

; FILING DATE: 12-Mar-2002

```

; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2268 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 69...2009
; OTHER INFORMATION:
; SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-10-096-241-31

```

```

Query Match          47.7%; Score 427.8; DB 14; Length 2268;
Best Local Similarity 89.8%; Pred. No. 1.6e-110;
Matches 459; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```

```

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
        || || | || | | | | || | || | | || |
Db      98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
        |||||
Db      158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
        |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 602
        |||||
Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 337

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
        |||||
Db      338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
        |||||
Db      398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy      723 CACCCTGTCATCCTGGTCTGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
        |||||
Db      458 CACCCTGTCATCCTGGTCTGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517

```


QY 783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842
 |||
 Db 518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577
 QY 843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873
 |||
 Db 578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608

RESULT 6

US-10-096-241-3

; Sequence 3, Application US/10096241

; Publication No. US20020127594A1

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
 ; AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/096,241

; FILING DATE: 12-Mar-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/699,591

; FILING DATE: 19-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Fasse, J. Peter

; REGISTRATION NUMBER: 32,983

; REFERENCE/DOCKET NUMBER: 07334/022001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-542-5070

; TELEFAX: 617-542-8906

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1607 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: Coding Sequence

; LOCATION: 79...621


```

; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/096,241
; FILING DATE: 12-Mar-2002
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: <Unknown>
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2467 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
;
; MOLECULE TYPE: cDNA
;
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 79...1893
;
; OTHER INFORMATION:
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-096-241-1

```

```

Query Match          45.2%; Score 405.4; DB 14; Length 2467;
Best Local Similarity 87.9%; Pred. No. 3.3e-104;
Matches 442; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

```

```

Qy      371 CCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 430
        | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2   CTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 61

Qy      431 GGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGA 490
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      62  GGCCCAAGCTGAAGAAGATGAAGAGCCAGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCA 121

Qy      491 AGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGG 550
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      122 AGTGTGAGGCAGCGCGGAAACCCCAAGCCTTCCTATCGCTGGTTCAAGGATGGCAAGG 181

Qy      551 AGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGAC 610
        | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

Qy      424 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 483
          |||
Db      27  GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 86

Qy      484 TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 543
          |||
Db      87  TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 146

Qy      544 GGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAG 596
          |||
Db      147 GGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAG 199

```

RESULT 9

US-10-029-386-12913

; Sequence 12913, Application US/10029386

; Publication No. US20030194704A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR GENE

; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

; FILE REFERENCE: AEOMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 12913

; LENGTH: 573

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO CHR5.3

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.55

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.72

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.66

; OTHER INFORMATION: SWISSPROT HIT: O14511, EVALUE 2.00e-28

; OTHER INFORMATION: NT HIT: AF119152.1, EVALUE 0.00e+00

; OTHER INFORMATION: EST_HUMAN HIT: BG996653.1, EVALUE 1.00e-108

US-10-029-386-12913

Query Match 19.3%; Score 173; DB 13; Length 573;

Best Local Similarity 100.0%; Pred. No. 8.9e-39;

Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      424 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 483
          |||
Db      377 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 436

Qy      484 TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 543
          |||
Db      437 TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 496

```

Qy 544 GGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAG 596
 |||
 Db 497 GGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAG 549

RESULT 10

US-10-029-386-2532/c

; Sequence 2532, Application US/10029386

; Publication No. US20030194704A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
 USEFUL FOR GENE

; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

; FILE REFERENCE: AEOMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 2532

; LENGTH: 579

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO CHR5.1

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.9

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.5

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.2

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.6

; OTHER INFORMATION: NT HIT: AF119153.1, EVALUE 0.00e+00

; OTHER INFORMATION: EST_HUMAN HIT: BF108794.1, EVALUE 2.00e-57

; OTHER INFORMATION: SWISSPROT HIT: O14511, EVALUE 2.00e-12

US-10-029-386-2532

Query Match 12.7%; Score 113.6; DB 13; Length 579;

Best Local Similarity 92.9%; Pred. No. 5.1e-22;

Matches 130; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

Qy 594 CAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGT 653
 |||
 Db 489 CAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGT 430

Qy 654 CTGCGAGGCCGAGAACATCCTGGGGAAGGACACCG-TCCGGGGCCGGCTTTACGTCAACA 712
 |||
 Db 429 CTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGCTCCGGGGCCGGCTTTACGTCAACA 370

Qy 713 GCGTGAGCACCACCCTGTCA 732
 ||| || ||| |||
 Db 369 GCGGTAGGTGGGCCCAGACA 350

RESULT 11


```
; APPLICANT: Ruben, Steven M.
; APPLICANT: Jonak, Zdenka L.
; APPLICANT: Trulli, Stephen H.
; APPLICANT: Fronwald, James A.
; APPLICANT: Terrett, Jonathan A.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1488.1070006
; CURRENT APPLICATION NUMBER: US/09/373,658
; CURRENT FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
; LENGTH: 1986
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:Unknown
US-09-373-658-71
```

```
Query Match          9.4%; Score 84; DB 11; Length 1986;
Best Local Similarity 49.0%; Pred. No. 1.6e-13;
Matches 361; Conservative 0; Mismatches 345; Indels 30; Gaps 4;
```

```
Qy      189 GGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGGGGGCTGCA 248
          ||  ||| || |  |  ||  || | |||  |||| |  | || | | |||
Db      759 GGTGTGGGCGGTGAAAGCCGGGGGCTTGAAGAAGGACTCGCTGCTCACCGTGCGCCTGGG 818

Qy      249 GCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAGCGCTACAT 308
          | | |  |  | | |  ||||| | |  |||| |  | | |  |||||
Db      819 GACCTGGGGCCACCCCGCCTTCCCCTCCTGCGGGAGGCTCAAGGAGGACAGCAGGTACAT 878

Qy      309 CTTTTTCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACG-----GCCTT 356
          ||| ||| ||||| || |  |  |  |  |  ||  ||||
Db      879 CTTCTTCATGGAGCCCAGCCTAACAGCACCAGCCGCGCGCCGGCCGCTTCCGAGCCTC 938

Qy      357 TGCCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCAC 416
          |  ||||  |  |||  ||  ||||| ||||  ||  |  |  |||||
Db      939 TTTCCCCCCTCTGGAGACGGGCCGGAACCTCAAGAAGGAGGTGAGCCGGGTGCTGTGCAA 998

Qy      417 TGAAGTGCAGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGA 476
          |||||  |  ||| |  ||||  ||||| |||||  |  |||
Db      999 GCGGTGCGCCTTGCCCTCCCAATTGAAAGAGATGAAAAGCCAGGAATCGGCTGCAGGTTTC 1058

Qy      477 GAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTT 536
          || | |  ||  ||||| |  |  |  |  |  |  |||
Db      1059 CAAACTAGTCCTTCGGTGTGAAACCAGTTCTGAATACTCCTCTCTCAGATTCAAGTGGTT 1118

Qy      537 CAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCA---TCAAATATGGCAACGG 593
          |||| |||| || ||  | || || |  |  |  |  |||  ||
Db      1119 CAAGAATGGGAATGAATTGAATCGAAAAACAAACCACAAAATATCAAGATACAAAAAA 1178

Qy      594 CAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGT 653
          |  ||| |||  ||| |  |  ||||| |  |||  ||  ||||  ||||| |
Db      1179 GCCAGGGAAGTCAGAACTTCGCATTAACAAAGCATCACTGGCTGATTCTGGAGAGTATAT 1238

Qy      654 CTGCGAGGCCGAGAACATCCTGGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAG 713
          ||| | |  |  ||  |  || ||  ||||  |  |  |  |  |  ||
```


Db 1239 GTGCAAAGTGATCAGCAAATTAGGAAATGACAGTGCCTCTGCCAATATCACCATCGTGGA 1298
 Qy 714 CGTGAGCACCACCCTGTCATCC-----TGGTCGGGGGCACGCCCCGGAAGTGCAACGAGAC 767
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1299 ATCAAACGCTACATCTACATCCACCACTGGGACAAGCCATCTTGTAATAATGTGCGGAGAA 1358
 Qy 768 AGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCT 827
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1359 GGAGAAAACCTTTCTGTGTGAATGGAGGGGAGTGCCTTCATGGTGAAAGACCTTTCAAACCC 1418
 Qy 828 CTC-----CTGCAAAGTGTCCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGC 878
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1419 CTCGAGATACTTGTGCAAAGTGCCCAAATGAGTTTACTGGTGATCGCTGCCAAACTACGT 1478
 Qy 879 AATGGTCAACTTCTCC 894
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1479 AATGGCCAGCTTCTAC 1494

RESULT 13

US-09-989-687-71

; Sequence 71, Application US/09989687

; Publication No. US20040002449A1

; GENERAL INFORMATION:

; APPLICANT: Hastings, Gregg A.

; APPLICANT: Ruben, Steven M.

; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides

; FILE REFERENCE: 1488.107000D

; CURRENT APPLICATION NUMBER: US/09/989,687

; CURRENT FILING DATE: 2001-11-21

; NUMBER OF SEQ ID NOS: 126

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 71

; LENGTH: 1986

; TYPE: DNA

; ORGANISM: Unknown

; FEATURE:

; OTHER INFORMATION: Description of Unknown Organism:Unknown

US-09-989-687-71

Query Match 9.4%; Score 84; DB 12; Length 1986;

Best Local Similarity 49.0%; Pred. No. 1.6e-13;

Matches 361; Conservative 0; Mismatches 345; Indels 30; Gaps 4;

Qy 189 GGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGCCGCTCCGGAGCGGGGGGCTGCA 248
 | | | | | | | | | | | | | | | | | | | | | |
 Db 759 GGTGTGGGCGGTGAAAGCCGGGGGCTTGAAGAAGGACTCGCTGCTCACCGTGCGCCTGGG 818
 Qy 249 GCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAGCGCTACAT 308
 | | | | | | | | | | | | | | | | | | | | | |
 Db 819 GACCTGGGGCCACCCCGCCTTCCCCTCCTGCGGGAGGCTCAAGGAGGACAGCAGGTACAT 878
 Qy 309 CTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACG-----GCCTT 356
 | | | | | | | | | | | | | | | | | | | | | |
 Db 879 CTTCTTCATGGAGCCCAGGCCAACAGCACCAGCCGCGCGCCGGCCGCTTCCGAGCCTC 938
 Qy 357 TGCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCAC 416

Db 939 TTTCCCCCTCTGGAGACGGGCCGGAACCTCAAGAAGGAGGTCAGCCGGGTGCTGTGCAA 998
 Qy 417 TGA CTGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGA 476
 Db 999 GCGGTGCGCCTTGCCTCCCCAATTGAAAGAGATGAAAAGCCAGGAATCGGCTGCAGGTTC 1058
 Qy 477 GAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCTACCGTTGGTT 536
 Db 1059 CAACTAGTCCTTCGGTGTGAAACCAGTTCTGAATACTCTCTCAGATTCAAGTGGTT 1118
 Qy 537 CAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCA---TCAAATATGGCAACGG 593
 Db 1119 CAAGAATGGGAATGAATTGAATCGAAAAACAAACCACAAAATATCAAGATACAAAAAA 1178
 Qy 594 CAGAAAGAACTCAGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGT 653
 Db 1179 GCCAGGGAAGTCAGAACTTCGCATTAACAAAGCATCACTGGCTGATTCTGGAGAGTATAT 1238
 Qy 654 CTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAG 713
 Db 1239 GTGCAAAGTGATCAGCAAATTAGGAAATGACAGTGCCTCTGCCAATATCACCATCGTGA 1298
 Qy 714 CGTGAGCACCACCCTGTCATCC-----TGGTCGGGGCACGCCCCGGAAGTGCAACGAGAC 767
 Db 1299 ATCAAACGCTACATCTACATCCACCACTGGGACAAGCCATCTTGTAATAATGTGCGGAGAA 1358
 Qy 768 AGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCT 827
 Db 1359 GGAGAAAACCTTCTGTGTGAATGGAGGGGAGTGCTTCATGGTGAAAGACCTTTCAAACCC 1418
 Qy 828 CTC-----CTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTTCGC 878
 Db 1419 CTCGAGATACTTGTGCAAGTGCCCAAATGAGTTTACTGGTGATCGCTGCCAAAACCTACGT 1478
 Qy 879 AATGGTCAACTTCTCC 894
 Db 1479 AATGGCCAGCTTCTAC 1494

RESULT 14

US-08-736-019-21

; Sequence 21, Application US/08736019

; Publication No. US20030207799A1

; GENERAL INFORMATION:

; APPLICANT: Goodearl, Andrew

; APPLICANT: Stroobant, Paul

; APPLICANT: Minghetti, Luisa

; APPLICANT: Waterfield, Michael

; APPLICANT: Marchionni, Mark

; APPLICANT: Chen, Mario

; APPLICANT: Hiles, Ian

; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR

; TITLE OF INVENTION: PREPARATION AND USE

; NUMBER OF SEQUENCES: 189

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Clark & Elbing LLP

```

; STREET: 176 Federal Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible Pentium
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/736,019
; FILING DATE: 22-OCT-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/471,833
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 91 07566.3
; FILING DATE: 10-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 04585/00200Q
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 428-0200
; TELEFAX: (617) 428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2003
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: N in positions 31 and 32 could be
; OTHER INFORMATION: either A or G.
US-08-736-019-21

```

```

Query Match          9.4%; Score 84; DB 7; Length 2003;
Best Local Similarity 49.0%; Pred. No. 1.6e-13;
Matches 361; Conservative 0; Mismatches 345; Indels 30; Gaps 4;

```

Qy 189 GGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGGGGGCTGCA 248
 || ||| || | | || || | ||| ||| | | | | |||
 Db 762 GGTGTGGGCGGTGAAAGCCGGGGGCTTGAAGAAGGACTCGCTGCTCACCGTGCGCCTGGG 821

Qy 249 GCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACAGCGCTACAT 308
 | | | | | | | | | | | | | | | | | | | | | |
 Db 822 GACCTGGGGCCACCCCGCCTTCCCCTCCTGCGGGAGGCTCAAGGAGGACAGCAGGTACAT 881

Qy 309 CTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACG-----GCCTT 356
 ||| ||| ||||| | | | | | | | | | | |||
 Db 882 CTTCTTCATGGAGCCCGACGCCAACAGCACCAGCCGCGCGCCGGCCGCTTCCGAGCCTC 941

Qy 357 TGCCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCAC 416
 | |||| | ||| || ||||| |||| | | |||||
 Db 942 TTTCCCCCTCTGGAGACGGGCGGAACCTCAAGAAGGAGGTGAGCCGGGTGCTGTGCAA 1001

Qy 417 TGAAGTGCAGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGA 476
 ||||| | ||| | |||| | |||| | |||| | |||
 Db 1002 GCGGTGCGCCTTGCTTCCCAATTGAAAGAGATGAAAGCCAGGAATCGGCTGCAGGTTC 1061

Qy 477 GAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGAATCCCCAGCCTTCTACCGTTGGTT 536
 || | | || ||||| | | | | | | | | |||||
 Db 1062 CAACTAGTCCTTCGGTGTGAAACCAGTTCTGAATACTCCTCTCTCAGATTCAAGTGGTT 1121

Qy 537 CAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGA---TCAAATATGGCAACGG 593
 |||| |||| || || | || || | | | | |||| ||
 Db 1122 CAAGAATGGGAATGAATTGAATCGAAAAACAAACCACAAAATATCAAGATACAAAAAA 1181

Qy 594 CAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGT 653
 | ||| ||| ||| ||| ||| | |||| | ||| || ||||| |||
 Db 1182 GCCAGGGAAGTCAGAACTTCGCATTAACAAAGCATCACTGGCTGATTCTGGAGAGTATAT 1241

Qy 654 CTGCGAGGCCGAGAATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAG 713
 ||| | | | || | || | || | ||| | | | | | |
 Db 1242 GTGCAAAGTGATCAGCAAATTAGGAAATGACAGTGCCTCTGCCAATATCACCATCGTGGA 1301

Qy 714 CGTGAGCACCACCCTGTCATCC-----TGGTCGGGGCACGCCCAGGAGTGAACGAGAC 767
 | | | | | |||| | || | || | || | ||||
 Db 1302 ATCAAACGCTACATCTACATCCACCACTGGGACAAGCCATCTTGTAATGTGCGGAGAA 1361

Qy 768 AGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCT 827
 | || | | || ||||| | |||| | | | | | |
 Db 1362 GGAGAAAACCTTCTGTGTGAATGGAGGGGAGTGCTTCATGGTGAAAGACCTTTCAAACCC 1421

Qy 828 CTC-----CTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTTCGC 878
 ||| ||||| || | | || || | || || | |||
 Db 1422 CTCGAGATACTTGTGCAAGTGCCCAAATGAGTTTACTGGTGATCGCTGCCAAAACCTACGT 1481

Qy 879 AATGGTCAACTTCTCC 894
 |||| | |||| |
 Db 1482 AATGGCCAGCTTCTAC 1497

; Publication No. US20030040465A1
; GENERAL INFORMATION:
; APPLICANT: Gywnne, David I.
; APPLICANT: Mahanthappa, Nagesh K.
; APPLICANT: Marchionni, Mark A.
; APPLICANT: Bermingham-McDonogh, Olivia
; APPLICANT: Goldin, Stanley M.
; APPLICANT: McBurney, Robert N.
; TITLE OF INVENTION: USE OF NEUREGULINS AS MODULATORS OF
; TITLE OF INVENTION: CELLULAR COMMUNICATION
; FILE REFERENCE: 04585/041005
; CURRENT APPLICATION NUMBER: US/09/366,886
; CURRENT FILING DATE: 1999-08-04
; PRIOR APPLICATION NUMBER: US 08/341,018
; PRIOR FILING DATE: 1994-11-17
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 2003
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (265)...(1530)
; NAME/KEY: variation
; LOCATION: (31)...(32)
; OTHER INFORMATION: n can be a or g.
US-09-366-886-71

Query Match 9.4%; Score 84; DB 11; Length 2003;
Best Local Similarity 49.0%; Pred. No. 1.6e-13;
Matches 361; Conservative 0; Mismatches 345; Indels 30; Gaps 4;

Qy	189	GGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGGGGGCTGCA	248
Db	762	GGTGTGGGCGGTGAAAGCCGGGGGCTTGAAGAAGGACTCGCTGCTCACCGTGC GCCTGGG	821
Qy	249	GCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAGCGCTACAT	308
Db	822	GACCTGGGGCCACCCCGCCTTCCCTCCTGCGGGAGGCTCAAGGAGGACAGCAGGTACAT	881
Qy	309	CTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACG-----GCCTT	356
Db	882	CTTCTTCATGGAGCCCAGCCAACAGCACCAGCCGCGCGCCGGCCGCCTTCCGAGCCTC	941
Qy	357	TGCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCAC	416
Db	942	TTTCCCCCTCTGGAGACGGGCCGAACCTCAAGAAGGAGGTGAGCCGGGTGCTGTGCAA	1001
Qy	417	TGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGA	476
Db	1002	GCGGTGCGCCTTGCCCTCCCAATTGAAAGAGATGAAAAGCCAGGAATCGGCTGCAGGTTT	1061
Qy	477	GAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTT	536
Db	1062	CAAAC TAGTCCTTCGGTGTGAAACCAGTTCTGAATACTCCTCTCTCAGATTCAAGTGGTT	1121

